

Research Paper

Population genetics of *Glossina palpalis palpalis* in sleeping sickness foci of Côte d'Ivoire before and after vector control

Djakaridja Berté^{a,b,1}, Thierry De Meeûs^{c,*,1}, Dramane Kaba^a, Modou Séré^d, Vincent Djohan^{a,b}, Fabrice Courtin^c, Martial N'Djetchi Kassi^e, Mathurin Koffi^e, Vincent Jamonneau^c, Bi Tra Dieudonné Ta^{a,b}, Philippe Solano^c, Eliezer Kouakou N'Goran^{b,2}, Sophie Ravel^{c,2}

^a Institut Pierre Richet/Institut National de Santé Publique, Bouaké, Côte d'Ivoire

^b Université Felix Houphouët-Boigny, Abidjan, Côte d'Ivoire

^c Intertryp, IRD, Cirad, Univ Montpellier, Montpellier, France

^d Université de Dédougou (UDDG), Dédougou, Burkina Faso

^e Laboratoire des Interactions Hôte-Microorganisme-Environnement et Evolution, Unité de Formation et de Recherche Environnement, Université Jean Lorougnon Guédé, Daloa, Côte d'Ivoire

ARTICLE INFO

Keywords:

Tsetse flies
Population genetics
Control
Dispersal
Resistance
Côte d'Ivoire

ABSTRACT

Glossina palpalis palpalis remains the major vector of sleeping sickness in Côte d'Ivoire. The disease is still active at low endemic levels in Bonon and Sinfra foci in the western-central part of the country. In this study, we investigated the impact of a control campaign on *G. p. palpalis* population structure in Bonon and Sinfra foci in order to adapt control strategies. Genetic variation at microsatellite loci was used to examine the population structure of different *G. p. palpalis* cohorts before and after control campaigns. Isolation by distance was observed in our sampling sites. Before control, effective population size was high (239 individuals) with dispersal at rather short distance (731 m per generation). We found some evidence that some of the flies captured after treatment come from surrounding sites, which increased the genetic variance. One Locus, GPCAG, displayed a 1000% increase of subdivision measure after control while other loci only exhibited a substantial increase in variance of subdivision. Our data suggested a possible trap avoidance behaviour in *G. p. palpalis*. It is important to take into account and better understand the possible reinvasion from neighboring sites and trap avoidance for the sake of sustainability of control campaigns effects.

1. Introduction

Human African Trypanosomiasis (HAT) is a centuries-old disease that has affected the lifestyle of people in sub-Saharan Africa (Steverding, 2008). It is a parasitic disease due to two subspecies of *Trypanosoma brucei* transmitted by tsetse flies belonging to the *Glossina* genus. There are two forms of HAT: one, known as gambiense HAT, due to *T. brucei gambiense*, is endemic in West and Central Africa and causes over 95% of current cases; the other, known as rhodesiense HAT, due to *T. brucei rhodesiense*, is endemic in East and Southern Africa and accounts for the remainder of cases (Büscher et al., 2017).

The disease reemerged at the end of the 1990s, but renewed efforts from endemic countries brought the disease under control again (Franco et al., 2018). In this context, sustainable elimination of the

gambiense HAT was considered as a feasible target for 2030 (Franco et al., 2014).

Tsetse control has recently become a key component of the overall activities of HAT control (Solano et al., 2013; Bouyer et al., 2015; Courtin et al., 2015). However, many tsetse control efforts were not sustainable due to either flies surviving the initial interventions, or flies immigrating from untreated regions, or both (Hargrove, 2003; Adam et al., 2014; Meeûs et al., 2019), except when control itself is sustained (Simo and Rayaisse, 2015; Meyer et al., 2016). This led to the fact that in 2015, animal trypanosomiasis was still an important issue in West and Central Africa (Simo and Rayaisse, 2015). The knowledge of the genetic structure of a target population can facilitate decision-making (Mccoy, 2008; Solano et al., 2010a; Solano et al., 2010b). Quantifying exchanges of individuals among subpopulations gives information on

* Corresponding author.

E-mail address: thierry.demeeus@ird.fr (T. De Meeûs).

¹ These authors contributed equally to this work.

² These authors supervised equally this work.

the isolation status and structure of the studied population. These population parameters are important for medical entomologists as they may have consequences on the epidemiology and control of vector borne diseases (De Meeûs et al., 2007; Solano et al., 2010b; Chevillon et al., 2012; Krafur and Maudlin, 2018; De Meeûs et al., 2019; Manangwa et al., 2019).

In Côte d'Ivoire, the development of cash crops (coffee and cocoa) in western part of the country has led to profound changes in the biotopes that are favorable to tsetse flies (Courtin et al., 2008; Cecchi et al., 2009) so that the country is the second most affected by HAT in West Africa (Simarro et al., 2010). Bonon and Sinfra foci are still active at low endemic level, since 11 cases have been reported in both foci in the last five years according to National Program for HAT Elimination. In order to accelerate elimination, in addition to active and passive HAT cases screening, a tsetse control campaign based on the use of impregnated tiny targets (Esterhuizen et al., 2011; Rayaisse et al., 2011; Rayaisse et al., 2012) has been firstly set in the Bonon focus in February 2016, then in Sinfra in May 2017.

In this study, we aimed to assess how vector control can affect the population genetic structure of *Glossina palpalis palpalis* in the Bonon focus. Sinfra subsamples were added to allow widening the studied area and get more precision on the initial structure of tsetse populations before any control in the Marahoué region.

2. Material and methods

2.1. Study site

The study was carried out in the Bonon and Sinfra HAT foci located in the western-central part of Côte d'Ivoire (Fig. 1). These foci belong to the Marahoué region, which is located in a mesophilic forest area, although forest has now almost disappeared, replaced by coffee and cocoa plantations. In these foci and surroundings, many livestock farms raising pigs, goats, sheep and cattle are present (Courtin et al., 2005; N'Djetchi et al., 2017). Due to the cash crops interests (coffee, cocoa, bananas, etc.) and associated human settlements, favoring tsetse-human contact, HAT found ideal geographic conditions for its development, and Bonon and Sinfra became the most active foci in Côte d'Ivoire at the beginning of the 1990s (Dje et al., 2002; Solano et al., 2003; Kaba et al., 2006). Thanks to control efforts, only few cases are now reported each year in these foci (Koffi et al., 2016).

2.2. Sampling

Sampling extended from May 2015 to March 2017. If we consider a 2 months length for a tsetse generation (Williams, 1990; Krafur, 2009), our sampling contain different tsetse fly cohorts: cohorts 1, 7, 8, 9, 10 and 11. In Bonon focus, the first entomological survey was carried out in June 2015 (T0, cohort 1), before the control campaign. Tsetse control campaign took place in February 2016 and then 30 sentinel traps were deployed at chosen sites in order to monitor the evolution of tsetse flies densities, among which 25 were used for the present study. Evaluations were done every three months allowing us to define T1, T2, T3 and T4. T1 (June 2016, cohort 7) corresponded to the first survey after control, T2 (September 2016, cohort 8) to the second survey, T3 (December 2016, cohort 10) to the third survey and T4 (March 2017, cohort 11) to the fourth entomological survey after control. In Sinfra, two entomological surveys were done before tsetse control. The first was done in May 2015 (T0, cohort 1) and the second in November–December 2016 (T0_bis, cohort 9). Sinfra subsamples allowed getting more precision on the initial structure of tsetse populations of the Marahoué region before control. Tsetse flies were sampled using Vavoua traps deployed during two consecutive days in tsetse fly favorable biotopes. In total, 17 and 8 traps were used in Bonon and Sinfra respectively (Table 1). For the population genetics analyses, we used lies captured in 14 traps: 6 in Bonon at T0 and 8 in Sinfra. In Bonon, we used captured

flies from 4 traps at T1, 5 traps at T2, 3 traps at T3, and 3 traps at T4. These 30 subsamples varied in size from 1 to 25 flies (13 on average), but most (21) contained > 9 flies. Surface of sampling was computed for each site (Table 1) with the longer distance between two traps (D_{max}) taken as the radius of the disc that contained all traps of the site ($S_s = \pi \times D_{max}^2$). Census densities (D_c) were computed as the number of captured flies per site (N_c) divided by S_s (Table 1). The comparison between densities at T0 and the others (T1–T4) was undertaken with a Wilcoxon rank sum test with continuity correction with R-Commander package (Fox, 2005; Fox, 2007) for R (R-Core-Team, 2018).

Three legs were removed from each fly and stored in 70% ethanol tubes labelled with a code containing the trap number followed by individual fly number and the sampling date. In total 403 tsetse individuals were analyzed (Table 1 and Supplementary Table S1). All flies were identified as *Glossina palpalis palpalis* according to morphological criteria (Pollock, 1982).

2.3. Genotyping

Tsetse flies were genotyped at 10 microsatellite loci (Supplementary Tables S1 and S2). Locus X55-3 is from (Solano et al., 1997). Loci XpGp13 and pGp24 come from (Luna et al., 2001) and GPCAG from (Baker and Krafur, 2001). Loci, B3, XB104, XB110 and C102 were kindly supplied by A. Robinson, Insect Pest Control Laboratory (formerly Entomology Unit), Food and Agricultural Organization of the United Nations/International Atomic Energy Agency [FAO/IAEA], Agriculture and Biotechnology Laboratories, Seibersdorf, Austria. Finally, loci pGp20 and pGp27 came from a microsatellite bank of *G. palpalis gambiense* (S. Ravel, personal communication). Those the name of which begins with the letter X are X-linked. We thus analyzed two data sets: one without these loci but with all individuals and the other with females only and all loci. More characteristics for describing the loci used can be found in the Supplementary Table S2.

In the laboratory, legs were dried and then subjected to chelex treatment as previously described (Ravel et al., 2007). The 10 PCR reactions were then carried out in a thermocycler (MJ Research, Cambridge, UK) in 20 µl final volume, using 10 µl of the diluted supernatant from the extraction step as template. After PCR amplification at the microsatellite loci, allele bands were routinely resolved on ABI 3500 XL sequencer (Applied Biosystems, USA). This method allows multiplexing and the use of four dyes (blue, red, green and yellow). Allele calling was done using GeneMapper V 4.1 software (Applied Biosystems) and the size standard GS600LIZ short run.

In order to confirm subspecies determination, we amplified partial sequences of ITS1 following the protocol described by Dyer et al. (2008).

All data and genotypes are available in the Supplementary Table S1. In total, 403 flies, including 309 females and 94 males, were genotyped.

2.4. Data analyses

All data were formatted for Create (Coombs et al., 2008) and transformed in the appropriate format for subsequent analyses.

Significance of linkage disequilibrium (LD) between locus pairs was assessed with the *G*-based test of Fstat 2.9.4 (Goudet, 2003) an updated version of Fstat (Goudet, 1995) with 10,000 permutations. The *G*-based test allows obtaining a global test across subsamples (for each pair of loci) and is more powerful than other combining procedures (De Meeûs et al., 2009). There are as many tests as locus pairs, i.e. $L(L-1)/2$ (15 if the number of loci $L = 6$) These tests are not independent. We used the Benjamini and Yekutieli false discovery rate (FDR) procedure (Benjamini and Yekutieli, 2001) that is appropriate in case of non-independent test series. The corrected *p*-values were computed with R (R-Core-Team, 2018).

For a hierarchy with three levels (individuals in subsample in total sample), three *F*-statistics (Wright, 1965) can be defined: F_{IS} , which

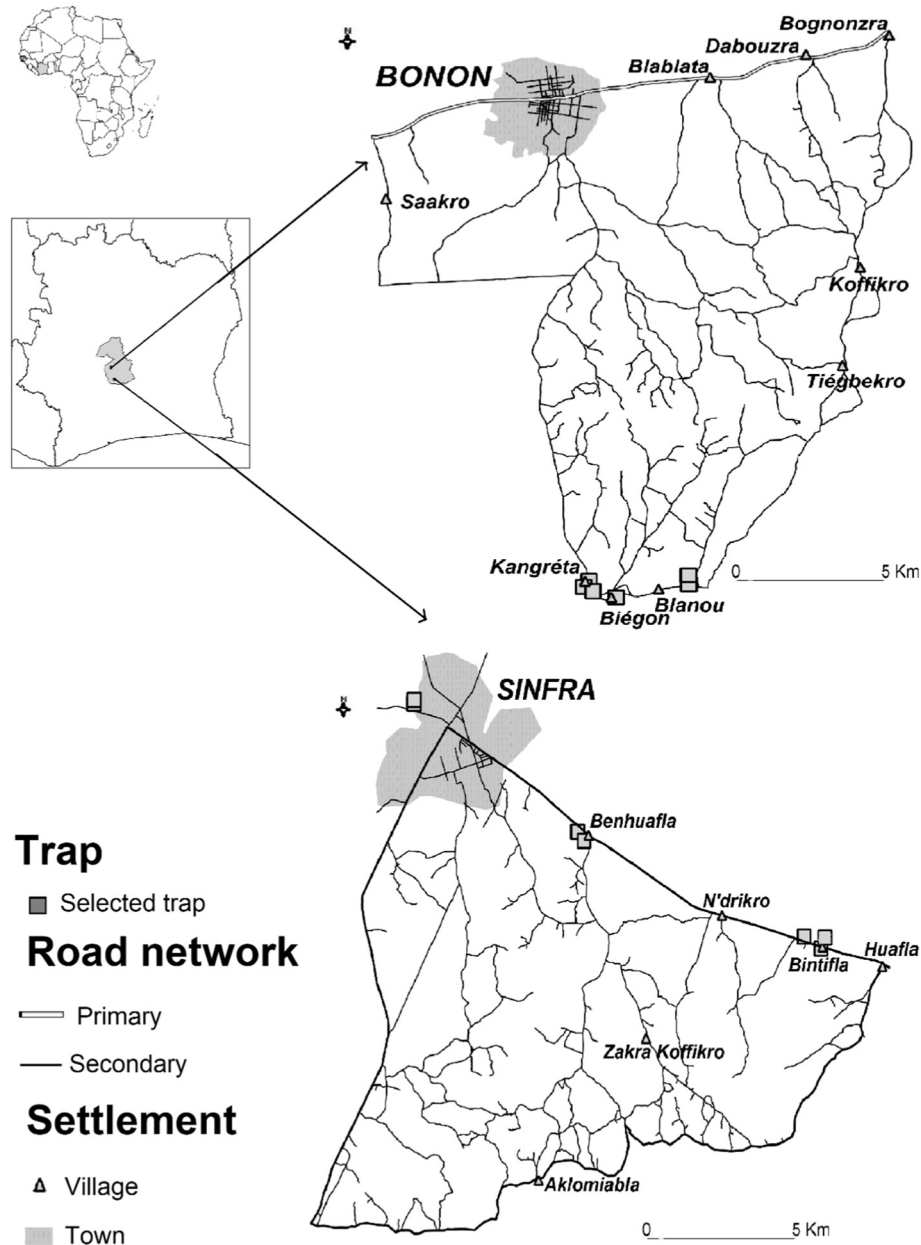


Fig. 1. Location map of study area with selected traps.

measures inbreeding of individuals relative to inbreeding of subsamples; F_{ST} , which measures inbreeding of subsamples relative to total inbreeding; and F_{IT} , which measures inbreeding of individuals relative to total inbreeding. Deviations of genotypic proportions expected under local panmixia are measured by F_{IS} , while F_{ST} measures the effect of subdivision (genetic isolation between subsamples) and F_{IT} reflects the combination of both (e.g. (De Meeûs et al., 2007)). Under the null hypothesis (panmixia and no subdivision), all these statistics are expected to be null. Otherwise F_{IS} and F_{IT} can vary from -1 (heterozygote excess) to $+1$ (homozygote excess) and F_{ST} from 0 (all subsamples share similar allele frequencies) to $+1$ (all subsamples fixed for one or the other allele). In any case, the three statistics are linked by the famous relationship: $1 - F_{IT} = (1 - F_{IS})(1 - F_{ST})$ (e.g. (De Meeûs, 2018)).

In dioecious species (like tsetse flies), heterozygote excess is expected over all loci in small random-mating subpopulations (e.g. (De Meeûs et al., 2007)). Multilocus positive F_{IS} (homozygote excess) can be produced by systematic mating between related individuals like sib-mating (e.g. (De Meeûs et al., 2007)). It can also come from the

admixture, in each subsample, of individuals that belong to genetically divergent entities (subpopulations, subspecies or species) (Wahlund effect) (e.g. (De Meeûs et al., 2007; De Meeûs, 2018)).

Technical problems, like null alleles, stuttering, short allele dominance or allele dropouts unevenly affects some loci, producing a positive F_{IS} with an important variation across loci (De Meeûs, 2018).

A positive value for F_{ST} suggests that the total population is subdivided, for instance, into n subpopulations of effective size N_e and an immigration rate of m . In an Island model of migration (with no spatial structure) (Wright, 1951) at mutation-drift equilibrium, we expect $F_{ST} = 1/[4N_e(m + u) + 1]$, where u is the mutation rate of the locus (e.g. (De Meeûs et al., 2007)). If $u \ll m$, the number of immigrants $N_e m$ can be extracted as $N_e m = (1 - F_{ST})/(4F_{ST})$.

Wright's F -statistics were estimated through Weir and Cockerham's unbiased estimators (Weir and Cockerham, 1984). It is worthy of note that θ , F_{ST} estimator, can display negative values. This happens when subsamples share more similar genetic composition than would be expected if the different subsamples were randomly drawn from the same

Table 1

Number of *Glossina palpalis palpalis* sampled in each site N_c , description of sampling, surface of sampling (S_s) (in km^2), observed densities (D_c) and number of genotyped flies (N_G). Treatment status (T) are indicated and GPS coordinates (Long, Lat) are given in decimal degrees.

Focus	T	Site	Trap	Long	Lat	S_s	N_c	D_c	N_G
Bonon	T0	Biegon	BE4017	-6.0298	6.7966	2.9326	9	85	0
Bonon	T0	Biegon	BE4019	-6.0377	6.7930	2.9326	38	85	0
Bonon	T0	Biegon	BG4018	-6.0344	6.7962	2.9326	203	85	22
Bonon	T0	Blanou	BE4015	-6.0250	6.7933	1.3513	38	33	0
Bonon	T0	Blanou	BG4013	-6.0239	6.7979	1.3513	0	33	0
Bonon	T0	Blanou	BG4014	-6.0231	6.7989	1.3513	5	33	0
Bonon	T0	Blanou	BG4016	-6.0249	6.7975	1.3513	1	33	0
Bonon	T0	Dianou	BE4008	-6.0170	6.7996	2.1569	259	336	15
Bonon	T0	Dianou	BE4012	-6.0124	6.7962	2.1569	63	336	0
Bonon	T0	Dianou	BG4005	-6.0149	6.7989	2.1569	0	336	0
Bonon	T0	Dianou	BG4006	-6.0157	6.7997	2.1569	7	336	0
Bonon	T0	Dianou	BG4007	-6.0170	6.8007	2.1569	79	336	0
Bonon	T0	Dianou	BG4009	-6.0164	6.8018	2.1569	111	336	0
Bonon	T0	Dianou	BG4010	-6.0171	6.8020	2.1569	206	336	14
Bonon	T0	Kangreta	BE4021	-6.0422	6.7989	0.3256	53	1056	16
Bonon	T0	Kangreta	BE4022	-6.0399	6.7978	0.3256	190	1056	21
Bonon	T0	Kangreta	BG4020	-6.0409	6.8006	0.3256	101	1056	15
Bonon	T1	Biegon	BG4018	-6.0344	6.7962	2.9326	22	8	18
Bonon	T1	Dianou	BE4008	-6.0170	6.7996	2.1569	1	1	1
Bonon	T1	Dianou	BG4010	-6.0171	6.8020	2.1569	2	1	2
Bonon	T1	Kangreta	BE4021	-6.0422	6.7989	0.3256	1	126	0
Bonon	T1	Kangreta	BE4022	-6.0399	6.7978	0.3256	3	126	3
Bonon	T1	Kangreta	BG4020	-6.0409	6.8006	0.3256	37	126	7
Bonon	T2	Biegon	BG4018	-6.0344	6.7962	2.9326	20	7	18
Bonon	T2	Dianou	BE4008	-6.0170	6.7996	2.1569	7	6	5
Bonon	T2	Dianou	BG4010	-6.0171	6.8020	2.1569	6	6	6
Bonon	T2	Kangreta	BE4021	-6.0422	6.7989	0.3256	14	86	11
Bonon	T2	Kangreta	BE4022	-6.0399	6.7978	0.3256	1	86	0
Bonon	T2	Kangreta	BG4020	-6.0409	6.8006	0.3256	13	86	11
Bonon	T3	Biegon	BG4018	-6.0344	6.7962	2.9326	0	0	0
Bonon	T3	Dianou	BE4008	-6.0170	6.7996	2.1569	10	21	6
Bonon	T3	Dianou	BG4010	-6.0171	6.8020	2.1569	36	21	17
Bonon	T3	Kangreta	BE4021	-6.0422	6.7989	0.3256	39	141	24
Bonon	T3	Kangreta	BE4022	-6.0399	6.7978	0.3256	1	141	1
Bonon	T3	Kangreta	BG4020	-6.0409	6.8006	0.3256	6	141	6
Bonon	T4	Biegon	BG4018	-6.0344	6.7962	2.9326	8	3	7
Bonon	T4	Dianou	BE4008	-6.0170	6.7996	2.1569	17	9	15
Bonon	T4	Dianou	BG4010	-6.0171	6.8020	2.1569	2	9	2
Bonon	T4	Kangreta	BE4021	-6.0422	6.7989	0.3256	0	3	0
Bonon	T4	Kangreta	BE4022	-6.0399	6.7978	0.3256	0	3	0
Bonon	T4	Kangreta	BG4020	-6.0409	6.8006	0.3256	1	3	0
Sinfra	T0	Benhuafla	SE2002	-5.8812	6.5960	0.2813	154	874	22
Sinfra	T0	Benhuafla	SE2004	-5.8796	6.5938	0.2813	92	874	19
Sinfra	T0	Bintifla	SE1012	-5.8189	6.5685	1.1297	55	283	19
Sinfra	T0	Bintifla	SE1013	-5.8178	6.5710	1.1297	207	283	25
Sinfra	T0	Bintifla	S1023	-5.8232	6.5712	1.1297	20	283	16
Sinfra	T0	Bintifla	S1031	-5.8178	6.5710	1.1297	38	283	14
Sinfra	T0	Ville	S5030	-5.9232	6.6272	0.0351	27	2279	13
Sinfra	T0	Ville	S5031	-5.9233	6.6263	0.0351	53	2279	12

subpopulation, i.e. when alleles are more related between than within subsamples (Weir, 1996) page 175).

For F -statistics, significant departure from 0 was tested by randomizing alleles among individuals within subsample (deviation from local random mating test) or of individuals among subsamples within the total sample (population subdivision test) (10,000 permutations in each case). The p -value then corresponded to the number of times a statistic measured in randomized samples was as big as or bigger than the observed one (unilateral tests). For F_{IS} , the statistic used was f (Weir and Cockerham's F_{IS} estimator). To test for subdivision, we used the G -based test (Goudet et al., 1996) over all loci, which is the most powerful procedure when combining tests across loci (De Meeûs et al., 2009).

To compute 95% confidence intervals (95%CI) of F -statistics, we used the standard error of F_{IS} (StrdErrFIS) and F_{ST} (StrdErrFST) computed by jackknife over populations, and 5000 bootstraps over loci as described elsewhere (De Meeûs et al., 2007). Since jackknife's computation of 95%CI assumes a normal distribution of the parameters and because F -statistics do not follow such a distribution, these confidence intervals were only used to graphically visualize parameter variation

across subsamples and not for statistical comparisons. Bootstrap's 95%CI does not require that the data follow any distribution and is thus statistically valid.

In case of significant homozygote excess and linkage disequilibrium we have tried to discriminate demographic from technical causes with the determination key proposed by De Meeûs (2018). In case of null alleles, both F_{IS} and F_{ST} are augmented, StrdErrFIS is at least twice StrdErrFST (jackknives over loci), a positive correlation is expected between F_{IS} and F_{ST} as is expected a positive correlation between F_{IS} and the number of missing data (putative null homozygotes). If such correlations do not exist and if $\text{StrdErr}F_{IS} > \text{StrdErr}F_{ST}$, then a Wahlund effect better explains the data. The significance of correlations was tested with a unilateral ($\rho > 0$) Spearman's rank correlation test with R. The presence of null alleles was also looked for with MicroChecker v 2.2.3 (Van Oosterhout et al., 2004) and null allele frequencies estimated with Brookfield's second method (Brookfield, 1996). The adjustment between observed and expected numbers of missing data was tested with a unilateral exact binomial test under R with the alternative hypothesis: "there are not enough missing data as

expected if heterozygote deficits were entirely explained by null alleles under panmixia". MicroChecker also checks for stuttering and short allele dominance (SAD). Short allele dominance was also assessed with unilateral ($\rho < 0$) Spearman's rank correlation test between allele size and F_{IT} , which is more powerful than other alternatives (Manangwa et al., 2019). In case of SAD, a negative correlation is expected (unilateral tests).

For local population structure studies, we used all subsamples, while for population subdivision studies we kept only comparisons between subsamples from the same cohort.

When null alleles are present, unbiased estimation of subdivision and/or of isolation by distance were obtained with the ENA correction for F_{ST} estimates and the INA correction for Cavalli-Sforza and Edwards' chord distance (Cavalli-Sforza and Edwards, 1967) D_{CSE} computed with FreeNA (Chapuis and Estoup, 2007). For this, all missing genotypes were converted into homozygous individuals for allele 999 as recommended (Chapuis and Estoup, 2007).

Isolation by distance was undertaken with Rousset's approach (Rousset, 1997) where, in a two dimension framework Rousset's Index $F_{ST,R} = F_{ST}/(1-F_{ST})$ follows $F_{ST,R} = a + b \times \text{Ln}(D_G)$ where a is a constant (intercept), $\text{Ln}(D_G)$ is the natural logarithm of the geographic distance between two sites and b is the slope of the regression. Rousset also demonstrated that the product of the effective population density D_e by the average of the squared axial distance between reproducing adults and their parents $\overline{\sigma^2}$: $D_e \times \overline{\sigma^2} = 1/(4\pi b)$; that the neighborhood size is $Nb = 1/b$; and that the number of immigrants from neighboring sites at each generation is $N_e m = 1/(2\pi b)$, where N_e is the effective subpopulation size and m is the immigration rate. The parameter σ represents half the average parent offspring axial distance. A proxy for dispersal distances per generation δ can be obtained if N_e and average surface of a subpopulation (S) are known: $D_e = N_e/S$ and $\delta \approx 2 \times \sqrt{1/(4\pi b D_e)}$. For S , we took either the surface computed above (S_S). Another possibility is to use the threshold distance (D_T) between two sites for F_{ST} to become positive, hence $D_T = e^{-a/b}$. This new parameter was then considered as the distance between the center of two neighboring subpopulations and hence to their diameter. The corresponding surface was then $S_T = \pi \times (D_T/2)^2$.

Significance of isolation by distance was assessed through the 5000 bootstraps confidence intervals from FreeNA and with a Mantel test between D_{CSE} and $\text{Ln}(D_G)$ or $F_{ST,R}$, computed with the INA or ENA correction by FreeNA, as recommended for microsatellite markers (Séré et al., 2017). Because some subsamples are not contemporaneous and because we kept only contemporaneous pairs, the final matrices were not squared and were thus analyzed with the menu "Mantelize it" of Fstat. Since Fstat handles bilateral tests, we converted the resulting p -value into a unilateral one (correlation between geographic and genetic distance is positive) by halving the bilateral p -value in case of positive correlation.

Effective population sizes (N_e) were estimated through several methods and softwares. The heterozygote excess method (Balloux, 2004) could not be used due to the excessive presence of null alleles in almost all loci (see below). The Linkage disequilibrium method (Waples and Do, 2010; Peel et al., 2013) and the coancestry method (Nomura, 2008) were undertaken with NeEstimator v 2.1 (Do et al., 2014). For LD method, several threshold values are proposed for minimal allele frequencies to be used (above 0, 0.01, 0.02 and 0.05). We used the average across usable values obtained with all methods. We also used the inter and intra correlation method (Vitalis and Couvet, 2001b) with the software Estim (Vitalis and Couvet, 2001a) (available at <http://www.tde-meeus.fr/ProgMeeusGB.html>). We then computed the average, minimal and maximal (MiniMax range) effective population sizes across methods weighted by the number of usable values obtained in each case. Effective population size reflects the demography and other phenomena as reproductive strategy and/or historical perturbations and its estimation also varies across methods (Krafsur and Maudlin, 2018). We nevertheless expected a good correlation between N_e and the

census size (N_c) of the corresponding subpopulations (De Meeüs et al., 2019).

The surface occupied by a subpopulation was inferred with S_S and S_T described above.

Effective population densities could then be computed as $D_e = N_e/S$.

The effect of control campaigns on the population genetic structure of *G. p. palpalis* was assessed first with a principal component analysis (PCA) undertaken with PCA-Gen 1.2.1 (Goudet, 1999) for which the metrics of principal axes correspond to Nei's G_{ST} (Nei and Chesser, 1983). Significance of axes was tested according to the broken stick criterion (Frontier, 1976) and 10,000 permutations of individuals across subsamples. We also compared F_{ST} (corrected for null alleles) between T0 subsample pairs and pairs between T0 and TX (i.e. sampled after control campaign at times T1, T2, T3 and T4) with a Mann-Whitney-Wilcoxon U test with Rcmdr.

3. Results

3.1. Fly densities

In Bonon, fly densities decreased significantly after the control campaign ($< 10\%$ of its initial value) as show in Fig. 2 (p -value = .0269) (see also Table 1).

3.2. Local population genetics with autosomal loci only in the Marahoue region

There was only one pair of loci in significant LD (p -value = .0406) that did not stay significant with Benjamini and Yekutieli correction (p -value = 1). We thus safely concluded that all markers are statistically independent.

There was an important heterozygote deficit that varied considerably from one locus to the other (Fig. 3). Null alleles only explained partly these results with a StdErrFIS more than four times StdErrFST , a positive though not significant correlation between the number of observed missing genotypes and F_{IS} ($\rho = 0.1429$, p -value = .4014); and a positive though not significant correlation between F_{IS} and F_{ST} ($\rho = 0.4058$, p -value = .2123). No stuttering could be evidenced and a significant short allele dominance was detected for locus pGp27 (Fig. 3). This locus thus seems problematic.

3.3. Local population genetics with females only in the Marahoue region

There were four pairs of loci in significant LD (p -values $< .05$) that did not stay significant after Benjamini and Yekutieli correction (p -values $> .9$).

The variation of F_{IS} was very important across the 10 loci with a

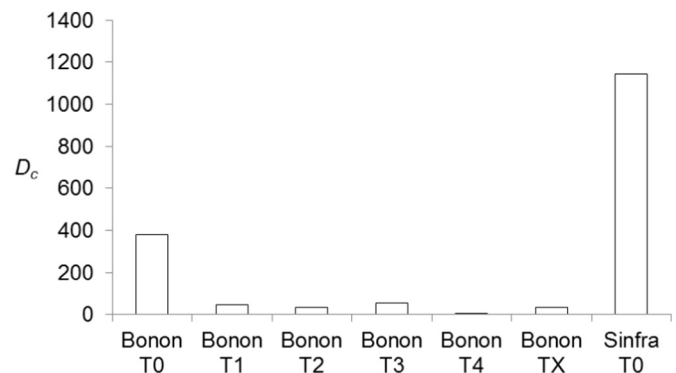


Fig. 2. Evolution of tsetse flies apparent density per km² (D_c) from T0 (before control campaign) in Bonon and Sinfra to T4 in Bonon. In Bonon, the difference between densities at T0 and TX ($X = 1-4$) was significant (p -value = .0269) (Wilcoxon rank sum test with continuity correction).

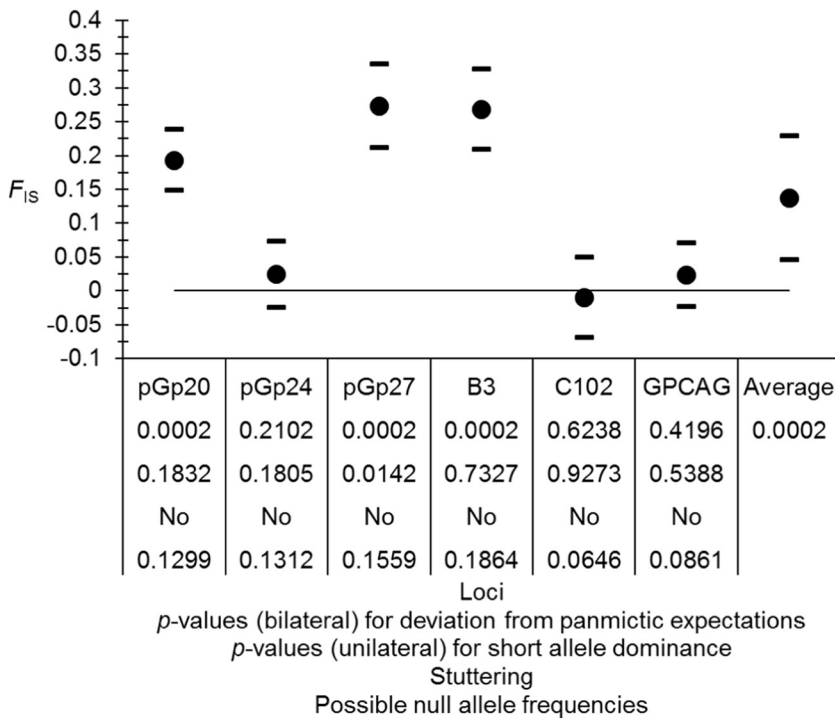


Fig. 3. F_{IS} observed in subsamples of *Glossina palpalis palpalis* from Ivory Coast for autosomal loci only with jackknife's over subsamples 95% confidence intervals for each locus and 95% Bootstrap over loci confidence interval for the average. Results of panmictic tests, short allele dominance tests, stuttering detection and possible null allele frequencies are also given.

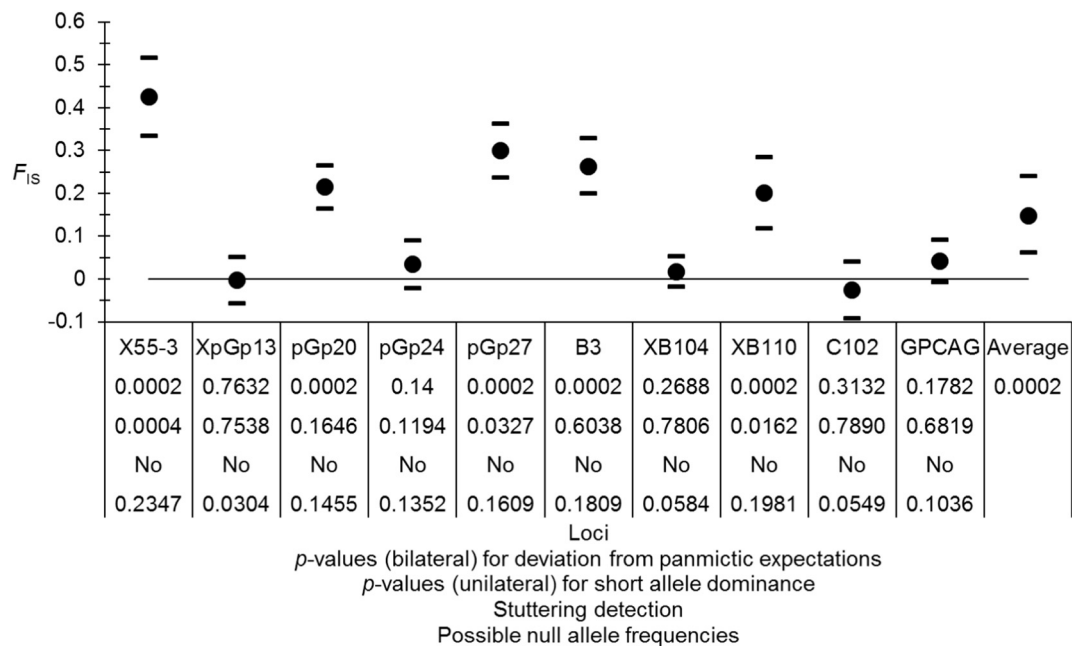


Fig. 4. F_{IS} observed in subsamples of females only of *Glossina palpalis palpalis* from Ivory Coast with jackknife's over subsamples 95% confidence intervals for each locus and 95% Bootstrap over loci confidence interval for the average. Results of panmictic tests, short allele dominance tests, stuttering detection and possible null allele frequencies are also given.

global highly significant heterozygote deficit (Fig. 4). There was some evidence of the presence of null alleles with a $StrdErrFIS$ eight times $StrdErrFST$, positive correlations between the number of missing data and F_{IS} ($\rho = 0.4681$) and between F_{IS} and F_{ST} ($\rho = 0.4667$) though marginally not significant (p -value = .0862 and p -value = .0891 respectively). Three loci displayed short allele dominance: Loci X55-3, pGp27 and XB110 (Fig. 4). This could be expected for the first two, which displayed the highest F_{IS} but with too few missing genotypes. Nevertheless, for XB110, the important number of missing genotypes (the highest observed) was in agreement with null alleles as the explanation for the important F_{IS} observed at this locus ($F_{IS} = 0.201$ with

26 missing genotypes and null allele frequency estimated as $p_n = 0.2$). Furthermore, it is worth noting that the SAD signature was mainly due to the fact that all longest alleles were rare (average frequency < 0.006) and displayed a very small F_{IT} (average < 0.042). Considering that with a frequency of $p_n = 0.2$, the probability for heterozygous individuals between such alleles and a null allele is $2 \times 0.2 \times 0.006 = 0.0024$ and the average of homozygote frequencies for these rare alleles is ~ 0.00005 , with a total sample size of 296, this means a total expected number of falsely interpreted or true homozygous individuals for these long and rare alleles is 0.725. Thus, because null alleles affect preferentially the F_{IS} of the most frequent

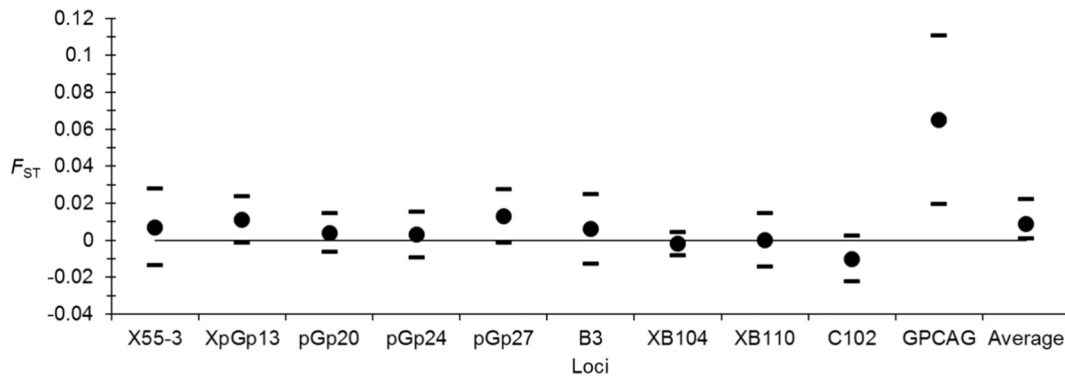


Fig. 5. Variation of F_{ST} across loci for females *Glossina palpalis palpalis* from Ivory Coast.

alleles, and because here these are also the shortest ones, this observation invalidates the SAD test for XB110. This locus can simply be considered as only affected by null alleles. Heterozygote deficits observed at other loci can be explained by null alleles. Subsequent subdivision measures and testing were thus done with corrections implemented by FreeNA.

Variations of F_{ST} across loci can be seen in the Fig. 5. From this figure, it can be seen that locus GPCAG's F_{ST} appeared completely outside the range of the nine other loci. Here, these subdivision measures are presented only for the sake of examining loci behavior. Indeed, values presented here both include temporal and spatial subdivision. Interestingly, this locus corresponds to a trinucleotide motive. It probably responds to some kind of selection.

In order to avoid possibly biased estimates, we removed loci X55-3, pGp27 and GPCAG from subsequent analyses.

To keep a number of loci above 5, we thus only studied female subsamples.

3.4. Population subdivision in the Marahoué region before control of female subsamples with the seven loci retained

The regression of isolation by distance is presented in the Fig. 6. The $F_{ST,R}$ based tests with 95% confidence interval of the slope or with the Mantel test are significant while the D_{CSE} based Mantel test is not. The 95% confidence interval of the slope is a very good indicator of what

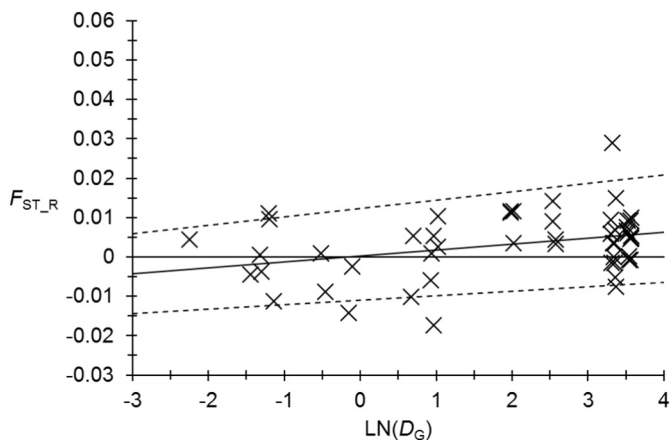


Fig. 6. Isolation by distance for contemporaneous subsamples of *Glossina palpalis palpalis* from Bonon and Sinfra in Ivory Coast before treatment (T0). The average regression line is represented as a straight line and the 95% confidence intervals (95%CI), obtained after 5000 bootstraps over loci, are in dotted lines. Abscissas are the natural logarithms of geographic distances in km. Ordinates are Rousset's $F_{ST,R}$ corrected for null alleles by FreeNA. Average slope was 0.0015 in 95%CI = [0.0011...0.0021]. Mantel test p -values were 0.4158 and 0.00795 for D_{CSE} and $F_{ST,R}$ respectively.

occurs and $F_{ST,R}$ based methods should be less powerful than with D_{CSE} (Séré et al., 2017), but it is highly significant here. It is thus safer to accept the alternative hypothesis of isolation by distance in this tsetse population.

Using the slope and its 95% confidence interval ($b = 0.0015$ in 95%CI = [0.0011...0.0021]), we could estimate a neighborhood size $N_b = 667$ individuals in 95%CI = [476...909] and an immigration of $N_e m = 106$ individuals from neighboring sites (in 95%CI = [76...145]).

Average effective subpopulation sizes was $N_e = 239$ ranging in MiniMax = [135...325]. Threshold geographic distance was $D_T = 875$ m, leading to a surface $S_T = 0.602$ km² while the average surface of sampling in sites was $S_S = 1.417$ km². Then effective population density $D_{e,S} = 169$ appeared much smaller than $D_{e,T} = 397$, which in turn displayed values that appeared closer to the average census density $D_c = 534$ individuals per km² (Fig. 7). Such inferences allowed computing dispersal as represented in the Fig. 8. Dispersal inferred from different effective population densities (Minimal, Average and Maximal) and methods (with threshold distance, sampling surface or census density) varied between 500 and 1800 m per generation. Dispersal obtained with sampling surfaces were significantly higher (1122 m on average in 95%CI [948...1310]) than those computed with the threshold surface (731 m in 95%CI = [618...854]), the last being very close to values calculated with the census population density (630 m in 95%CI = [533...636]).

3.5. Effects of control in Bonon

The PCA graphic is represented in the Fig. 9. Although no axis was significant, it can be seen that subsamples before control campaigns

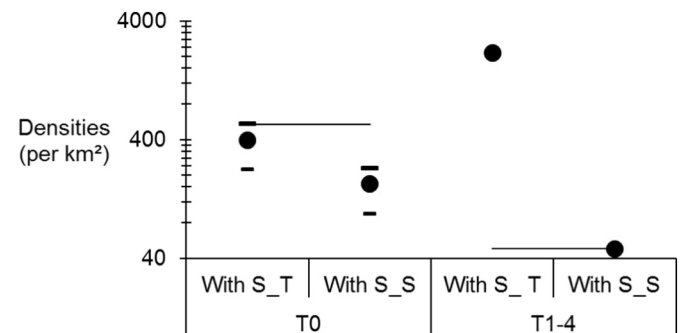


Fig. 7. Population densities of *Glossina palpalis palpalis* with three methods: effective population density using the threshold distance for genetic subdivision ($D_{e,T}$), effective subpopulation size using the average sampling surface of a site ($D_{e,S}$) and census density (D_c) (thin straight line). Effective population densities are represented with their average minimal and maximal values. Estimates are given for subsamples before the control campaign (T0) and after control (T1-4). Ordinates were scaled in log.

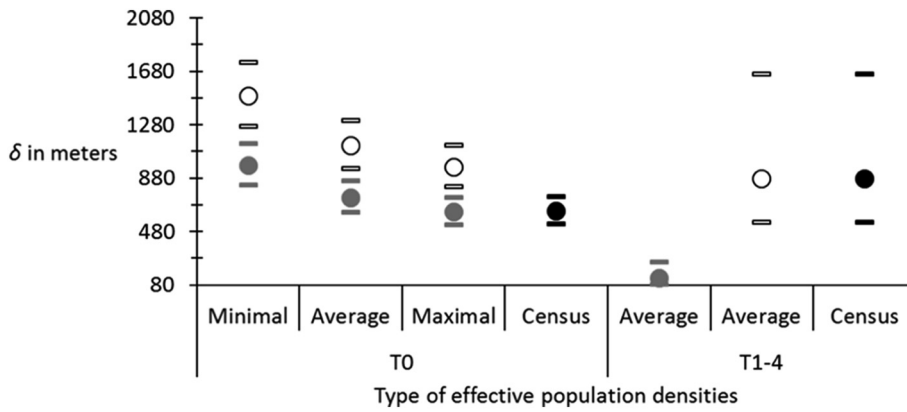


Fig. 8. Dispersal (δ) of *Glossina palpalis palpalis* in Bonon and Sinfra (Ivory Coast) for different effective population densities (minimal, averaged and maximal) and different methods: using the threshold distance for genetic subdivision for defining the surface occupied by a subpopulation (grey), using the average sampling surface of the different sites (empty symbols) and using the census population density. The 95% confidence intervals computed from the isolation by distance slope after 5000 bootstraps over loci are represented as small dashes. Estimates are given for subsamples before control campaigns (T0) and after (T1–4) and were all made from Rousset's isolation by distance model between contemporaneous subsample pairs only. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

occupy less space than subsamples after control campaigns, which suggests a greater genetic variance after control.

Accordingly, genetic subdivision appeared smaller between traps from T0 ($F_{ST} = -0.005$) than between traps at T0 and traps at TX ($F_{ST} = 0.0035$) but the difference was not significant ($p\text{-value} = .1225$).

Modelling isolation by distance between contemporaneous subsamples after control was uneasy because of the small number of available points. It nevertheless provided a slope the bootstrap confidence interval of which did not contain 0 ($b = 0.0086$ in 95%CI = [0.0024...0.0219]). From Fig. 7, effective population density significantly increases after treatment according to the distance threshold method ($D_{e,T}$). On the contrary, effective population density using sampling surface ($D_{e,S}$) significantly drops to values similar to the census density.

Regarding dispersal (Fig. 8), it significantly drops with the threshold distance method after the control campaign, while dispersal from the sampling surface method and from the census density are not significantly affected after the control campaign, with bootstrap 95%CI containing all the range of densities estimated at T0.

3.6. Comparison between the seven loci and GPCAG for T0/TX genetic differentiation in Bonon

As can be seen from Fig. 10, Locus GPCAG displayed a slightly higher subdivision measure as compared to the other loci for subsamples from T0 but displayed a 1000% increase when measured between T0 and TX, while other loci displayed an unchanged average though with a much higher variance. By examining population subdivision between T1–T4 subsamples revealed absence of genetic differentiation ($F_{ST_FreeNA} = -0.0127$, $p\text{-value} = .5843$). After pooling all

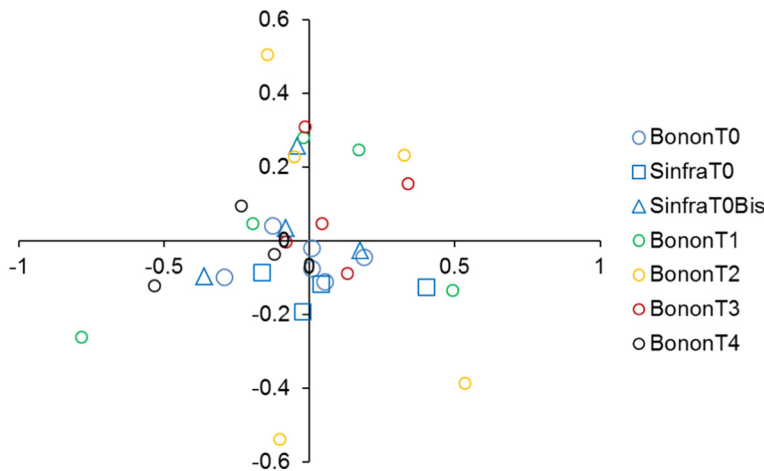


Fig. 9. PCA Analysis result on the two first principal axes for the different subsamples before control campaigns in Bonon and Sinfra (T0, in blue) and after control campaigns (T1, T2, T3, and T4, other colors). Axis percent inertia values are 11.68 and 8.08 for axis 1 and 2 respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

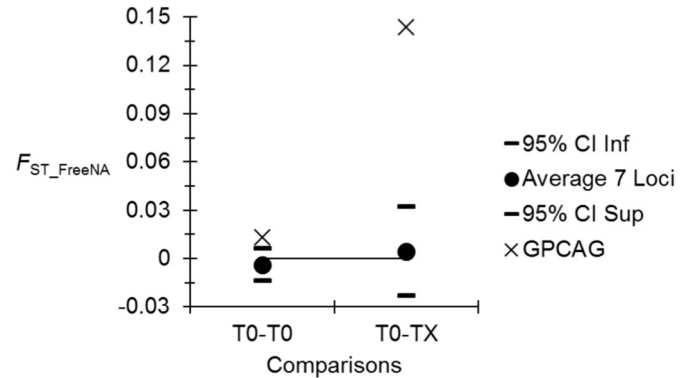


Fig. 10. Average subdivision measures corrected for null alleles with FreeNA (F_{ST_FreeNA}) between T0 subsample pairs (T0–T0) (before treatment campaigns) (black circles) with the seven loci kept for all analyses and 95% confidence intervals (dashes), for GPCAG alone (crosses) and between T0 and TX subsample pairs (T0–TX) ($X = 1, 2, 3$ or 4).

T0 traps together and all T1–4 together, we obtained a highly significant ($p\text{-value} < .0001$) subdivision ($F_{ST_FreeNA} = 0.1286$) for GPCAG consistent with what can be seen in Fig. 10. A glance at allele frequencies evolution reveals that this is mainly due to a 328% increase of allele 219 (from 0.146 to 0.625) at locus GPCAG after treatment.

4. Discussion

In the Marahoue region (Bonon and Sinfra), before control, we noted relatively high population densities and short dispersal distances

as compared to other studies (De Meeûs et al., 2019). Traps selected for this study were located in villages' edge with sacred forest (i.e. religiously protected from any human activity with economic purposes) that seems to constitute highly suitable environments for tsetse flies: ideal hygrometric and shade conditions, protection from insecticide pressure existing in cultivated fields and abundant hosts as wandering pigs (Laveissière et al., 1985; Sané et al., 2000a, 2000b). Flies may not need to disperse much because they get all they need there and/or, high densities limit settling possibilities of immigrants. Dispersal is indeed negatively density dependent in tsetse flies (De Meeûs et al., 2019).

In Bonon, the increase of genetic variance after control campaigns suggests that flies sampled after controls may come from surrounding sites, at least partly, as was observed 35 years ago for the same species in the same area (Randolph et al., 1984; Rogers and Randolph, 1984; Rogers et al., 1984). The drastic drop of > 91% in flies' density may have freed up space for surrounding flies. So the area that seemed saturated before control can receive new colonizers. Because of recent treatment and/or immigration, T1–4 subsamples were harvested from subpopulations in strong disequilibrium where effective population density and isolation by distance are hard to measure. The very odd results obtained for $D_{e,T}$ and δ_T after treatment can illustrate this. Alternatively, it may also reflect the inappropriateness of defining the surface occupied by a subpopulation with the threshold distance for genetic subdivision. The relevance of such method was indeed strongly questioned (Rousset, 1997). On the other hand, before control campaigns, these $D_{e,T}$ and δ_T appeared very close to values using census population sizes. When we used the sampling surface, effective population densities appeared significantly smaller than census density before control but dropped to values similar to those computed with census population sizes, while dispersal remained unchanged but with an important increase of the confidence interval. This may reflect a classical result where $N_e < N_c$ (e.g. (Krafsur and Maudlin, 2018)), before treatment, while both N_e and N_c drop to similar very low values after control and only the variance of dispersal is affected but not the average. So, estimating effective population density and dispersal with the sampling surface method may be much more accurate than the threshold distance for genetic subdivision method.

Locus GpCAG seems selected by the control technique used in Bonon, by a mechanism that may include either insecticide resistance, behavioral avoidance of trapping device or another unknown mechanism. This would mean that some flies with specific GpCAG profiles (allele 219) were able to escape from control measures: either because this trinucleotide locus may code for something, or because it is part of a selected gene as an intronic sequence, or very close to it. It would be interesting to identify what caused this result. We have been unable to find to what corresponded the sequence where this locus is exactly. Indeed, a “highly similar sequences (megablast)” in GenBank at <https://blast.ncbi.nlm.nih.gov/Blast.cgi> (with the accession number AY033512.1) outputted no result but the sequence deposited by Baker and Krafsur (2001). A discontinuous megablast in GenBank, using the Program BLASTN 2.8.1+ (Zhang et al., 2000) did not provide very useful information. It matched at 75% with a nuclear receptor coactivator of *Salvelinus alpinus* (a salmonid fish), with a putative mediator of RNA polymerase II transcription subunit 26 of *Ceratitis capitata* (at 69%) and *Rhagoletis zephyria* (at 72%), two Dipteran insects of the sub-order Brachycera like tsetse flies. Deeper investigations would be required to decipher the mechanism involved in the possible resistance associated with GPCAG allele 219. Insecticide resistance has never been reported in tsetse flies to our knowledge and we could find nothing more than recommendations on that issue in an old FAO report (Georghiou et al., 1993) and assumptions that the likelihood of insecticide resistance evolution is negligible in tsetse (Krafsur and Maudlin, 2018). Moreover, the slight superiority of subdivision measured at GPCAG locus even before the control campaign presented here suggests the signature of preceding past treatments. This can come from the different sampling campaigns that have preceded this study between 2000 and 2001

(Courtin et al., 2005; Ravel et al., 2007), using Vavoua traps without insecticide, which resulted in drastic reduction in tsetse densities (Courtin et al., 2005). Alternatively, the massive use of insecticides and pesticides in cocoa and coffee plantation may also explain these results, hypothetically. Resistance evidenced here is possibly behavioral. Insecticide-avoidance behavior is known for different mosquito species (Chareonviriyaphap et al., 1997; Chareonviriyaphap et al., 2013; Tainchum et al., 2013; Porciani et al., 2017). As for tsetse flies, physiological or behavioral resistance is only suspected, if not speculated (Georghiou et al., 1993) but, to our knowledge, was never documented.

If resistance is behavioral, the fact that a substantial proportion of resistant tsetse flies harboring the GpCAG-219 allele were captured in the trapping device after treatment suggests: i) that resistance is not absolute but statistical; and ii) that the proportion of resistant tsetse flies must be much higher in the pool of uncaptured flies.

More researches are needed to locate this locus in the *Glossina* genomes and check if it is in or near a coding gene, determine its nature, and clarify if this trinucleotide microsatellite itself is responsible for what we have evidenced in the present paper.

5. Conclusion

The work presented here shows that control campaigns has modified tsetse flies population structure. Although it has allowed reducing considerably tsetse fly's densities, it may also have selected for the emergence of flies resisting the treatment by a mechanism that remains to be identified. This result should be taken into account and new strategies developed to prevent reinvasion.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.meegid.2019.103963>.

Acknowledgments

This study was funded by the Education and Research Ministry of Côte d'Ivoire, as part of “Contrat Désendettement Développement” (C2D) managed by IRD (Institut de Recherche pour le Développement). The work presented is also part of “Targeting tsetse: a demonstrating project” and “Trypa-NO!”, which are supported by Bill & Melinda Gates Foundation. Djakaridja BERTE is supported by a grant from IRD, “Allocation de Recherche pour une Thèse au Sud”, Contrat No. 862286D. This research is part of the International Joint Laboratory on Vector Borne Disease (LAMIVECT) financed by the IRD.

References

- Adam, Y., Bouyer, J., Dayo, G.K., Mahama, C.I., Vreysen, M.J.B., Cecchi, G., Abd-Alla, A.M.M., Solano, P., Ravel, S., De Meeûs, T., 2014. Genetic comparison of *Glossina tachinoides* populations in three river basins of the upper west region of Ghana and implications for tsetse control. *Infect. Genet. Evol.* 28, 588–595.
- Baker, M.D., Krafsur, E.S., 2001. Identification and properties of microsatellite markers in tsetse flies *Glossina morsitans* sensu lato (Diptera: Glossinidae). *Mol. Ecol. Notes* 1, 234–236.
- Balloux, F., 2004. Heterozygote excess in small populations and the heterozygote-excess effective population size. *Evolution* 58, 1891–1900.
- Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188.
- Bouyer, J., Dicko, A.H., Cecchi, G., Ravel, S., Guerrini, L., Solano, P., Vreysen, M.J.B., De Meeûs, T., Lancelot, R., 2015. Mapping landscape friction to locate isolated tsetse populations candidate for elimination. *Proc. Natl. Acad. Sci. U. S. A.* 112, 14575–14580.
- Brookfield, J.F.Y., 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol. Ecol.* 5, 453–455.
- Büscher, P., Cecchi, G., Jamonneau, V., Priotto, G., 2017. Human African trypanosomiasis. *Lancet* 390, 2397–2409.
- Cavalli-Sforza, L.L., Edwards, A.W.F., 1967. Phylogenetic analysis: model and estimation procedures. *Am. J. Hum. Genet.* 19, 233–257.
- Cecchi, G., Courtin, F., Paone, M., Diarra, A., Franco, J.R., Mattioli, R.C., Simarro, P.P., 2009. Mapping sleeping sickness in Western Africa in a context of demographic transition and climate change. *Parasite* 16, 99–106.
- Chapuis, M.P., Estoup, A., 2007. Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* 24, 621–631.
- Chareonviriyaphap, T., Roberts, D.R., Andre, R.G., Harlan, H.J., Manguin, S., Bangs, M.J.,

1997. Pesticide avoidance behavior in *Anopheles albimanus*, a malaria vector in the Americas. *J. Am. Mosq. Control Assoc.* 13, 171–183.
- Chareonviriyaphap, T., Bangs, M.J., Suwonkerd, W., Kongmee, M., Corbel, V., Ngoen-Klan, R., 2013. Review of insecticide resistance and behavioral avoidance of vectors of human diseases in Thailand. *Parasit. Vect.* 6.
- Chevillon, C., McCoy, K.D., De Meeùs, T., 2012. Population genetics and molecular epidemiology of infectious diseases. In: Morand, S., Beaudreau, F., Cabaret, J. (Eds.), *New Frontiers of Molecular Epidemiology of Infectious Diseases*. Springer, Dordrecht, pp. 45–76.
- Coombes, J.A., Letcher, B.H., Nislow, K.H., 2008. CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. *Mol. Ecol. Res.* 8, 578–580.
- Courtin, F., Dupont, S., Zeze, D.G., Jamonneau, V., Sane, B., Coulibaly, B., Cuny, G., Solano, P., 2005. Human African trypanosomiasis: urban transmission in the focus of Bonon (Ivory Coast). *Tropical Med. Int. Health* 10, 340–346.
- Courtin, F., Jamonneau, V., Duvallet, G., Garcia, A., Coulibaly, B., Doumenge, J.P., Cuny, G., Solano, P., 2008. Sleeping sickness in West Africa (1906–2006): changes in spatial repartition and lessons from the past. *Tropical Med. Int. Health* 13, 334–344.
- Courtin, F., Camara, M., Rayaisse, J.B., Kagbadouno, M., Dama, E., Camara, O., Traore, I.S., Fouamba, J., Peylhard, M., Somda, M.B., Leno, M., Lehane, M.J., Torr, S.J., Solano, P., Jamonneau, V., Bucheton, B., 2015. Reducing human-tsetse contact significantly enhances the efficacy of sleeping sickness active screening campaigns: a promising result in the context of elimination. *PLoS Negl. Trop. Dis.* 9.
- De Meeùs, T., 2018. Revisiting F_{IS} , F_{ST} , Wahlund effects and null alleles. *J. Hered.* 109, 446–456.
- De Meeùs, T., McCoy, K.D., Prugnolle, F., Chevillon, C., Durand, P., Hurtrez-Boussès, S., Renaud, F., 2007. Population genetics and molecular epidemiology or how to “débusquer la bête”. *Infect. Genet. Evol.* 7, 308–332.
- De Meeùs, T., Guégan, J.F., Teriokhin, A.T., 2009. MultiTest V.1.2, a program to binomially combine independent tests and performance comparison with other related methods on proportional data. *BMC Bioinform.* 10, 443.
- Dje, N.N., Miezian, T.W., N'Guessan, P., Brika, P., Doua, F., Boa, F., 2002. Geographic distribution of trypanosomiasis treated in Ivory Coast from 1993 to 2000. *Bull. Soc. Path. Ex.* 95, 359–361.
- Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol. Ecol. Res.* 14, 209–214.
- Dyer, N.A., Lawton, S.P., Ravel, S., Choi, K.S., Lehane, M.J., Robinson, A.S., Okedi, L.M., Hall, M.J.R., Solano, P., Donnelly, M.J., 2008. 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.* 49, 227–239.
- Esterhuizen, J., Rayaisse, J.B., Tirados, I., Mpiana, S., Solano, P., Vale, G.A., Lehane, M.J., Torr, S.J., 2011. Improving the cost-effectiveness of visual devices for the control of riverine tsetse flies, the major vectors of human African trypanosomiasis. *PLoS Negl. Trop. Dis.* 5.
- Fox, J., 2005. The R commander: a basic statistics graphical user interface to R. *J. Stat. Software* 14, 1–42.
- Fox, J., 2007. Extending the R commander by “plug in” packages. *R News* 7, 46–52.
- Franco, J.R., Simarro, P.P., Diarra, A., Ruiz-Postigo, J.A., Jannin, J.G., 2014. The journey towards elimination of gambiense human African trypanosomiasis: not far, nor easy. *Parasitology* 141, 748–760.
- Franco, J.R., Cecchi, G., Priotto, G., Paone, M., Diarra, A., Grout, L., Simarro, P.P., Zhao, W., Argaw, D., 2018. Monitoring the elimination of human African trypanosomiasis: update to 2016. *PLoS Negl. Trop. Dis.* 12, e0006890.
- Frontier, S., 1976. Etude de la décroissance des valeurs propres dans une analyse en composantes principales: comparaison avec le modèle du bâton brisé. *J. Exp. Mar. Biol. Ecol.* 25, 67–75.
- Georghiou, G.I., Denholm, I., Heckel, D.G., Hemingway, J., Mouches, C., 1993. FAO/IAEA Consultants Group Meeting on “the Potential for Tsetse Flies to Develop Resistance to Insecticides”. Available at: http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/43/033/43033092.pdf.
- Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered.* 86, 485–486.
- Goudet, J., 1999. PCA-GEN for Windows. Available from: <http://www2.unil.ch/popgen/softwares/pcagen.htm>.
- Goudet, J., 2003. Fstat (ver. 2.9.4), a program to estimate and test population genetics parameters. Available at: <http://www.t-de-meeus.fr/Programs/Fstat294.zip> (Updated from Goudet (1995)).
- Goudet, J., Raymond, M., De Meeùs, T., Rousset, F., 1996. Testing differentiation in diploid populations. *Genetics* 144, 1933–1940.
- Hargrove, J.W., 2003. In: DFID Animal Health Programme (Ed.), *Tsetse Eradication: Sufficiency, Necessity and Desirability*. Centre for Tropical Veterinary Medicine, University of Edinburgh.
- Kaba, D., Dje, N.N., Courtin, F., Oke, E., Koffi, M., Garcia, A., Jamonneau, V., Solano, P., 2006. The impact of war on the evolution of sleeping sickness in west-central Cote d'Ivoire. *Tropical Med. Int. Health* 11, 136–143.
- Koffi, M., N'Djetchi, M., Ilboudo, H., Kaba, D., Coulibaly, B., N'Gouan, E., Kouakou, L., Bucheton, B., Solano, P., Courtin, F., Ehrhardt, S., Jamonneau, V., 2016. A targeted door-to-door strategy for sleeping sickness detection in low-prevalence settings in Cote d'Ivoire. *Parasite* 23, 51.
- Krafsur, E.S., 2009. Tsetse flies: genetics, evolution, and role as vectors. *Infect. Genet. Evol.* 9, 124–141.
- Krafsur, E.S., Maudlin, I., 2018. Tsetse fly evolution, genetics and the trypanosomiasis - a review. *Infect. Genet. Evol.* 64, 185–206.
- Laveissière, C., Couret, D., Staak, C., Hervouët, J.P., 1985. *Glossina palpalis* et ses hôtes en secteur forestier de Côte d'Ivoire: relation avec l'épidémiologie de la trypanosomiase humaine. *Ch. O.R.S.T.O.M., Sér. Ent. Méd. et Parasitol.* 23, pp. 297–303.
- Luna, C., Bonizzoni, M.B., Cheng, Q., Aksoy, S., Zheng, L., 2001. Microsatellite polymorphism in the tsetse flies (Diptera: Glossinidae). *J. Med. Entomol.* 38, 376–381.
- Manangwa, O., De Meeùs, T., Grébaut, P., Segard, A., Byamungu, M., Ravel, S., 2019. Detecting Wahlund effects together with amplification problems: cryptic species, null alleles and short allele dominance in *Glossina pallidipes* populations from Tanzania. *Mol. Ecol. Res.* 19, 757–772.
- Mccoy, K.D., 2008. The population genetic structure of vectors and our understanding of disease epidemiology. *Parasite* 15, 444–448.
- De Meeùs, T., Ravel, S., Solano, P., Bouyer, J., 2019. Negative density dependent dispersal in tsetse flies: a risk for control campaigns? *Trends Parasitol.* (In press).
- Meyer, A., Holt, H.R., Selby, R., Guitián, J., 2016. Past and ongoing tsetse and animal trypanosomiasis control operations in five African countries: a systematic review. *PLoS Negl. Trop. Dis.* 10, e0005247.
- N'Djetchi, M.K., Ilboudo, H., Koffi, M., Kabore, J., Kabore, J.W., Kaba, D., Courtin, F., Coulibaly, B., Fauret, P., Kouakou, L.A., Ravel, S., Deborggraeve, S., Solano, P., De Meeus, T., Bucheton, B., Jamonneau, V., 2017. The study of trypanosome species circulating in domestic animals in two human African trypanosomiasis foci of Cote d'Ivoire identifies pigs and cattle as potential reservoirs of *Trypanosoma brucei gambiense*. *PLoS Negl. Trop. Dis.* 11.
- Nei, M., Chesser, R.K., 1983. Estimation of fixation indices and gene diversities. *Ann. Hum. Genet.* 47, 253–259.
- Nomura, T., 2008. Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evol. Appl.* 1, 462–474.
- Peel, D., Waples, R.S., Macbeth, G.M., Do, C., Ovenden, J.R., 2013. Accounting for missing data in the estimation of contemporary genetic effective population size (N_e). *Mol. Ecol. Res.* 13, 243–253.
- Pollock, J.N., 1982. *Manuel de Lutte Contre la Mouche Tsé-tsé*, Volume 1: Biologie, Systématique et Répartition des Tsé-tsé. Food and Agriculture Organization of the United Nations (FAO), Rome.
- Porciani, A., Diop, M., Moiroux, N., Kadoko-Lambi, T., Cohuet, A., Chandre, F., Dormont, L., Penetier, C., 2017. Influence of pyrethroid-treated bed net on host seeking behavior of *Anopheles gambiae* s. s. carrying the kdr allele. *PLoS One* 12.
- Randolph, S.E., Rogers, D.J., Kuzoe, F.A.S., 1984. Local variation in the population dynamics of *Glossina palpalis palpalis* (Robineau-Desvoidy) (Diptera: Glossinidae). II. The effect of insecticidal spray programmes. *Bull. Entomol. Res.* 74, 425–438.
- Ravel, S., De Meeùs, T., Dujardin, J.P., Zeze, D.G., Gouling, R.H., Dufour, I., Sane, B., Cuny, G., Solano, P., 2007. The tsetse fly *Glossina palpalis palpalis* is composed of several genetically differentiated small populations in the sleeping sickness focus of Bonon, Côte d'Ivoire. *Infect. Genet. Evol.* 7, 116–125.
- Rayaisse, J.B., Esterhuizen, J., Tirados, I., Kaba, D., Salou, E., Diarrassouba, A., Vale, G.A., Lehane, M.J., Torr, S.J., Solano, P., 2011. Towards an optimal design of target for tsetse control: comparisons of novel targets for the control of *Palpalis* group tsetse in West Africa. *PLoS Negl. Trop. Dis.* 5.
- Rayaisse, J.B., Kröber, T., McMullin, A., Solano, P., Mihok, S., Guerin, P.M., 2012. Standardizing visual control devices for tsetse flies: West African species *Glossina tachinoides*, *G. palpalis gambiense* and *G. morsitans submorsitans*. *PLoS Negl. Trop. Dis.* 6, e1491.
- R-Core-Team, 2018. *R: A Language and Environment for Statistical Computing*, Version 3.5.0 (2018-04-23) Ed. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Rogers, D.J., Randolph, S.E., 1984. A review of density-dependent processes in tsetse populations. *Insect Sci. Appl.* 5, 397–402.
- Rogers, D.J., Randolph, S.E., Kuzoe, F.A.S., 1984. Local variation in the population dynamics of *Glossina palpalis palpalis* (Robineau-Desvoidy) (Diptera: Glossinidae). I. Natural population regulation. *Bull. Entomol. Res.* 74, 403–423.
- Rousset, F., 1997. Genetic differentiation and estimation of gene flow from F -statistics under isolation by distance. *Genetics* 145, 1219–1228.
- Sané, B., Laveissière, C., Méda, H.A., 2000a. Diversity of the diet of *Glossina palpalis palpalis* in the forest zone of Côte d'Ivoire: relation to the prevalence of African human trypanosomiasis. *Tropical Med. Int. Health* 5, 73–78.
- Sané, B., Laveissière, C., Méda, H.A., 2000b. Spatial distribution and bloodmeal preferences of *Glossina palpalis palpalis* in the forest focus of Zoukougbeu: epidemiological consequences. *Parasite* 7, 241–244.
- Séré, M., Thévenon, S., Belem, A.M.G., De Meeùs, T., 2017. Comparison of different genetic distances to test isolation by distance between populations. *Heredity* 119, 55–63.
- Simarro, P.P., Cecchi, G., Paone, M., Franco, J.R., Diarra, A., Ruiz, J.A., Fevre, E.M., Courtin, F., Mattioli, R.C., Jannin, J.G., 2010. The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *Int. J. Health Geogr.* 9.
- Simo, G., Rayaisse, J.B., 2015. Challenges facing the elimination of sleeping sickness in west and central Africa: sustainable control of animal trypanosomiasis as an indispensable approach to achieve the goal. *Parasit. Vect.* 8, 640.
- Solano, P., Duvallet, G., Dumas, V., Cuisance, D., Cuny, G., 1997. Microsatellite markers for genetic population studies in *Glossina palpalis* (Diptera: Glossinidae). *Acta Trop.* 65, 175–180.
- Solano, P., Kone, A., Garcia, A., SANE, B., Michel, V., Michel, J.F., Coulibaly, B., Jamonneau, V., Kaba, D., Dupont, S., FOURNET, F., 2003. Représentation spatiale des déplacements des malades dans un foyer de trypanosomose humaine africaine de Côte d'Ivoire. *Med. Trop.* 63, 577–582.
- Solano, P., Kaba, D., Ravel, S., Dyer, N.A., Sall, B., Vreysen, M.J., Seck, M.T., Darbyshire, H., Gardes, L., Donnelly, M.J., De Meeus, T., Bouyer, J., 2010a. Population genetics as a tool to select tsetse control strategies: suppression or eradication of *Glossina palpalis gambiense* in the Niayes of Senegal. *PLoS Negl. Trop. Dis.* 4, e692.

- Solano, P., Ravel, S., De Meeûs, T., 2010b. How can tsetse population genetics contribute to African trypanosomiasis control? *Trends Parasitol.* 26, 255–263.
- Solano, P., Torr, S.J., Lehane, M.J., 2013. Is vector control needed to eliminate gambiense human African trypanosomiasis? *Front. Cell Infect. Microbiol.* 3, 33.
- Steverding, D., 2008. The history of African trypanosomiasis. *Parasit. Vect.* 1, 3.
- Tainchum, K., Polsomboon, S., Grieco, J.P., Suwonkerd, W., Prabaripai, A., Sungvornyothin, S., Chareonviriyaphap, T., Achee, N.L., 2013. Comparison of *Aedes aegypti* (Diptera: Culicidae) resting behavior on two fabric types under consideration for insecticide treatment in a push-pull strategy. *J. Med. Entomol.* 50, 59–68.
- Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P., 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4, 535–538.
- Vitalis, R., Couvet, D., 2001a. ESTIM 1.0: a computer program to infer population parameters from one- and two-locus gene identity probabilities. *Mol. Ecol. Notes* 1, 354–356.
- Vitalis, R., Couvet, D., 2001b. Estimation of effective population size and migration rate from one- and two-locus identity measures. *Genetics* 157, 911–925.
- Waples, R.S., Do, C., 2010. Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol. Appl.* 3, 244–262.
- Weir, B.S., 1996. *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*. Sinauer Associates, Sunderland, Massachusetts.
- Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- Williams, B.G., 1990. Tsetse fly (Diptera: Glossinidae) population dynamics and the estimation of mortality rates from life-table data. *Bull. Entomol. Res.* 80, 479–485.
- Wright, S., 1951. The genetical structure of populations. *Ann. Eugenics* 15, 323–354.
- Wright, S., 1965. The interpretation of population structure by F-statistics with special regard to system of mating. *Evolution* 19, 395–420.
- Zhang, Z., Schwartz, S., Wagner, L., Miller, W., 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214.