

The tsetse fly *Glossina palpalis palpalis* is composed of several genetically differentiated small populations in the sleeping sickness focus of Bonon, Côte d'Ivoire

S. Ravel^a, T. de Meeus^b, J.P. Dujardin^b, D.G. Zézé^{c,e}, R.H. Gooding^d,
I. Dusfour^a, B. Sané^e, G. Cuny^a, P. Solano^{a,e,*}

^aIRD UR 177, Laboratoire de Recherche et de Coordination sur les Trypanosomoses IRD/CIRAD,
Campus de Baillarguet, 34398 Montpellier Cedex 5, France

^bGEMI, UMR CNRS IRD 2724, Centre IRD Montpellier, 911 Avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France

^cUFR Sciences et Gestion de l'Environnement, Université Abobo Adjamé, 02 BP 801 Abidjan 02, Cote d'Ivoire

^dDepartment of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

^eInstitut Pierre Richet Bouaké, s/c INSP Abidjan, BP V47 Abidjan, Cote d'Ivoire

Received 30 March 2006; received in revised form 29 June 2006; accepted 3 July 2006

Available online 4 August 2006

Abstract

Glossina palpalis is the main vector of human African trypanosomiasis (HAT, or sleeping sickness) that dramatically affects human health in sub-Saharan Africa. Because of the implications of genetic structuring of vector populations for the design and efficacy of control campaigns, *G. palpalis palpalis* in the most active focus of sleeping sickness in Côte d'Ivoire was studied to determine whether this taxon is genetically structured. High and statistically significant levels of within population heterozygote deficiencies were found at each of the five microsatellite loci in two temporally separated samples. Neither null alleles, short allele dominance, nor trap locations could fully explain these deviations from random mating, but a clustering within each of the two samples into different genetic sub-populations (Wahlund effect) was strongly suggested. These different genetic groups, which could display differences in infection rates and trypanosome identity, were composed of small numbers of individuals that were captured together, leading to the observed Wahlund effect. Implications of this population structure on tsetse control are discussed.

© 2006 Elsevier B.V. All rights reserved.

Keywords: *Glossina palpalis*; Sleeping sickness; Côte d'Ivoire; Wahlund effect; *T. brucei gambiense*; Microsatellites; Population structure

1. Introduction

Tsetse flies (Diptera: Glossinidae) are the main vectors of trypanosomes (Kinetoplastida: Trypanosomatidae), which cause human and animal trypanosomiasis in tropical Africa. These diseases, after years of neglect, are strongly re-emerging and currently have a considerable impact on public health and economic development in sub-Saharan Africa (WHO, 2001; Louis, 2001), although there are finally recent signs of declining following a considerable intervention effort based on case treatment and surveillance (Jannin, 2005).

As with all vector-borne diseases, trypanosomiasis transmission is complex and requires at least three interacting organisms: the vertebrate host, the insect vector, and the pathogen. Interruption at any point in these interactions can potentially reduce disease transmission. Unfortunately, it has not been possible to develop vaccines against trypanosomes, because African trypanosomes frequently change the antigenic nature of their surface proteins. Current control efforts rely primarily on active surveillance and treatment. These efforts, however, are hampered by population mobility, poor sensitivity of diagnostic tests, and lack of low-cost, efficacious drugs that have minimal adverse side effects.

Since trypanosome transmission relies on tsetse flies, eliminating this vector appears to be the most effective strategy to break the disease cycle. This presupposes a thorough knowledge of the epidemiology of the disease, including the

* Corresponding author. Present address: IRD 01 BP 171, Bobo Dioulasso 01, Burkina Faso. Tel.: +226 76 40 56 17; fax: +226 20 97 09 42.

E-mail address: solano@mpl.ird.fr (P. Solano).

biology, ecology, and population genetics of the targeted tsetse species and subspecies. However, our knowledge of tsetse still suffers from many gaps. For instance, it is widely acknowledged that genetic variation among vector populations probably affects the transmission of many parasitic diseases at a macrogeographic level (Lanzaro and Warburg, 1995; Gooding, 1996; Awono-Ambene et al., 2004), but little information is available to assess the spatial scale of this phenomenon. Tsetse flies have been reported to have a short flight duration (totalling 15–30 min/day), and may move, on average, from 200 m to 1.4 km/day (Leak, 1999), this mobility being very dependent on microhabitat conditions, in particular humidity. However, compared to this high daily mobility, Cuisance et al. (1985) reported that most *Glossina palpalis gambiensis* have a net lifetime displacement of no more than 1 or 2 km (in a humid

savannah area); the exceptions are some old females that can disperse more than 5 km. Data are scarce for forest species, but Gouteux et al. (1983) reported lifetime dispersal of a few kilometres for *Glossina palpalis palpalis*.

Glossina palpalis sensu lato is the most important vector of human African trypanosomiasis (HAT or sleeping sickness) in West Africa, and is also one of the most important vectors of animal trypanosomoses. Two subspecies of *G. palpalis* are generally recognized: *G. p. gambiensis* and *G. p. palpalis*. The former lives in humid savannah, and the latter in forested areas; their distribution limits follow more or less the savannah–forest transition (Challier et al., 1983; see also Fig. 1). Preliminary work on *G. p. gambiensis* in Burkina Faso using microsatellite markers detected a genetic structuring of populations that may have an impact on the ability of this subspecies to transmit

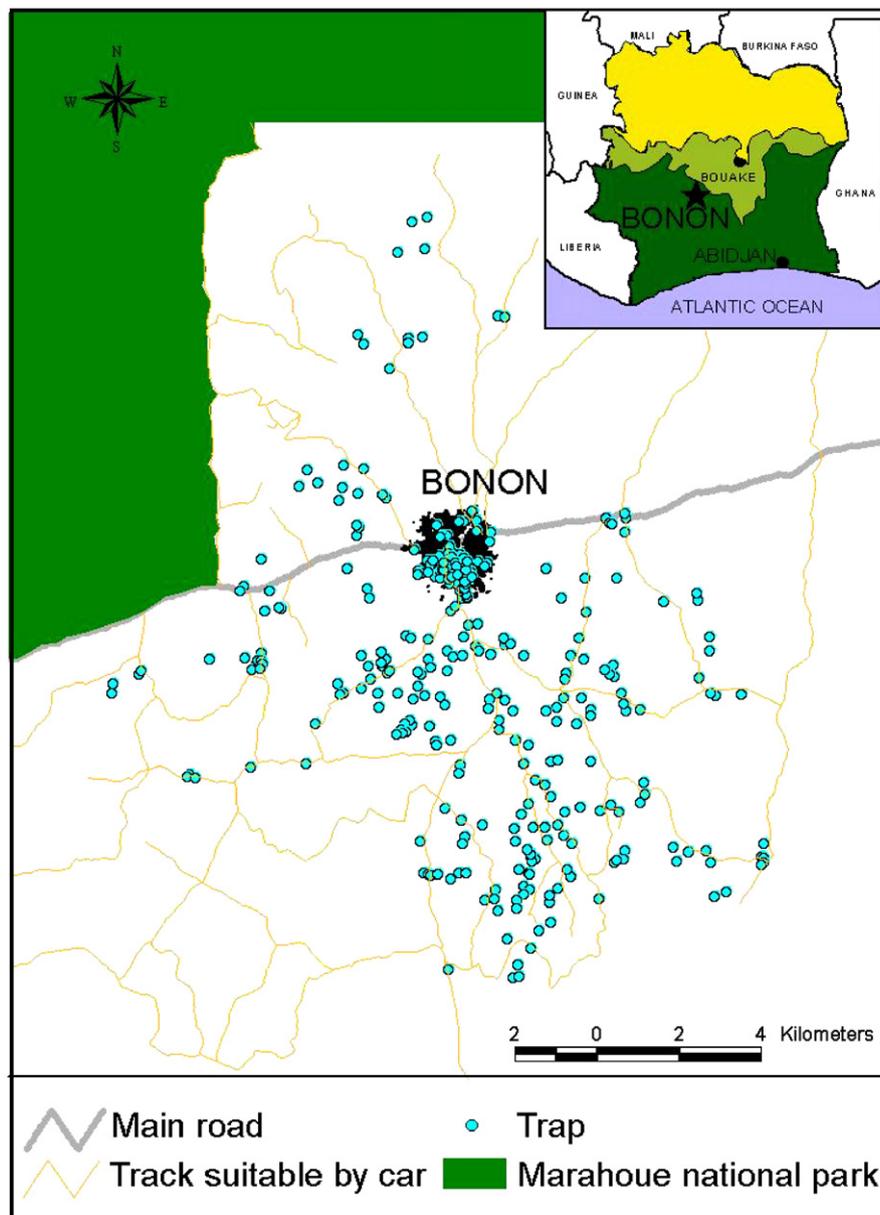


Fig. 1. Geographic location of the study area and the traps. The sleeping sickness focus of Bonon is located in the Center West of Côte d'Ivoire, just under the savannah–forest transition zone, and a few kilometres south of the reported limit of the two subspecies of *Glossina palpalis* (Challier et al., 1983).

animal trypanosomes (Solano et al., 2000). The objective of the present work was to investigate population genetics of *G. p. palpalis*, the main vector of HAT in West African forest zones, and its possible consequences for the epidemiology and control of this disease in an active sleeping sickness focus in west-central Côte d'Ivoire.

2. Materials and methods

2.1. Study area

The HAT focus of Bonon is located in west-central Côte d'Ivoire (6°55'N–6°W). Although the area is located in the forest area, near the forest–savannah transition (see Fig. 1), the forest has nearly been replaced by cocoa/coffee plantations and other crops. This new agro-ecosystem favours *G. palpalis* over other tsetse species because of the broad host preferences of *G. palpalis* (Gouteux et al., 1982; Reid et al., 2000). A few kilometres to the north the savannah area begins, and there *G. p. gambiense* is found (Challier et al., 1983).

Beginning in April 2000, an HAT survey was conducted (Solano et al., 2003). Between 2000 and 2003, 130 cases of HAT were detected in this focus, giving an infection rate of about 0.45% (Kaba et al., 2006), which makes the area of Bonon the most active HAT focus of the country. After the first medical survey of 2000, two entomological surveys were conducted in November 2000 and January 2001. These tsetse samples will be referred to in the text as sample 1 and sample 2 for November 2000 and January 2001, respectively.

2.2. Entomological survey, tsetse dissection, sample collection

Following the first medical survey which detected 74 patients (Solano et al., 2003), the key epidemiologically significant sites (home, water supply points, and working places) relevant to each patient were recorded by Global Positioning System (GPS). Subsequently 320 Vavoua traps (Laveissière and Grébaud, 1990) were placed in those areas frequented by the patients, which represent a total surface area of about 15 km × 15 km. Each trap was maintained for four days, with cages changed daily, and the number of tsetse and sex ratio recorded.

At the field camp, after removing the tsetse from the trap, individual *G. p. palpalis* were processed as follows:

Three legs were removed and stored in dry Eppendorf tubes for subsequent genetic analyses.

Age of females was determined by dissection of the ovaries (Challier, 1965).

The mouthparts, salivary glands, and midgut were dissected to detect trypanosome infections by microscopy.

When a tsetse was found infected with trypanosomes, each of the three organs was collected separately in Eppendorf tubes containing 50 µl sterile distilled water. This material was used for PCR identification of trypanosomes using taxon-specific DNA primers (Masiga et al., 1992; Majiwa

et al., 1994; Solano et al., 1995), including a microsatellite marker specific for *T. b. gambiense* group 1 (Truc et al., 2002). The details of techniques and results regarding identification of trypanosomes by PCR in the tsetse have been reported in Jamonneau et al. (2004).

2.3. PCR reactions at microsatellite loci

DNA was extracted from 122 females and 2 males from sample 1 (originating from 57 out of the 320 traps), and 92 females and 9 males from sample 2 (originating from 47 out of the 320 traps) as follows. To each tube containing the legs of the tsetse, 200 µl of 5% Chelex[®] chelating resin was added (Walsh et al., 1991; Solano et al., 2000). After incubation at 56 °C for 1 h, DNA was denatured at 95 °C for 30 min. The tubes were then centrifuged at 12,000 × *g* for 2 min and frozen for later analysis.

The PCR reactions were carried out in a thermocycler (MJ Research, Cambridge, UK) in 50 µl final volume, using 10 µl of the supernatant from the extraction step. After PCR amplification, allele bands were routinely resolved on 10% non-denaturing polyacrylamide gels stained with ethidium bromide. To determine the different allele sizes, a large panel of about 30 size markers was used. These size markers had been previously generated by cloning alleles from individual tsetse flies into pGEM-T Easy Vector (Promega Corporation, Madison, WI, USA). Three clones of each allele were sequenced using the T7 primer and the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, Foster City, CA, USA). Sequences were analysed on a PE Applied Biosystems 310 automatic DNA sequencer (PE Applied Biosystems) and the exact size of each cloned allele was determined. PCR products from these cloned alleles were run in the same acrylamide gel as the samples, allowing the allele size of the samples to be determined accurately.

Five microsatellite loci were used for genetic studies: *Gpg55.3* (Solano et al., 1997) which had been isolated from *G. p. gambiense*, and *Pgp24*, *Pgp13*, *Pgp11*, and *Pgp1*, isolated from a laboratory colony of *G. p. palpalis* (Luna et al., 2001). At the time of this study these were the only loci available for tsetse, and they had been shown to be highly polymorphic and to show easily read banding patterns in laboratory-reared tsetse flies and in those from a preliminary sample from the field population being studied (data not shown).

The microsatellite locus *Gpg55.3* is on the X-chromosome (Solano et al., 2000; Gooding et al., 2004). Because of this and the predominance of females in the samples, we analysed mostly field-collected females.

2.4. Analysis of microsatellite data

For each population, Wright's F_{is} (within sample heterozygote deficiency, a measure of deviation from panmixia) and F_{st} (measure of population differentiation) were estimated using Weir and Cockerham's (1984) unbiased estimators (f for F_{is} , θ for F_{st}). These estimators were calculated with FSTAT V. 2.9.3 software (Goudet, 1995). For random mating (within

samples) or random distribution of individuals (between samples), F values are expected to be zero.

The significance of F_{is} (deviation from panmixia) was also tested for each locus, and for all loci, in each sample, and simultaneously in several samples, using 10,000 permutations of alleles between individuals in FSTAT. Males were hemizygous at locus *Gpg55.3*. For that locus, measure of F_{is} and its significance were conducted only on females. The significance of F_{st} (population differentiation) was assessed using 10,000 permutations of genotypes among samples (FSTAT). Linkage disequilibrium was tested by the exact test of GENEPOP 3.3 for genotypic linkage disequilibrium (Raymond and Rousset, 1995).

The frequency of null alleles that was necessary to explain our observed F_{is} was calculated following Brookfield's (1996) method. The comparison between the expected occurrence of null individuals (blanks) and the real blanks observed was undertaken by a unilateral exact binomial test under S-Plus 2000 professional release 1. Short allele dominance can be recognised by regressing the F_{is} of each allele against allele sizes (see Wattier et al., 1998; De Meelis et al., 2002), we thus attempted to find such a pattern in the available data using S-Plus 2000 professional release 3. The basic assumption of the model is that in a heterozygous individual, the shortest allele will be better (or faster) amplified by a factor proportional to the size difference between the two alleles in concern. Then a Spearman correlation coefficient and its P -value were computed, together with R^2 (regression's coefficient of determination), this latter measuring the percentage of the F_{is} explained by allele size. As the relationship expected under the short allele dominance hypothesis is negative, we only kept the regressions giving such a relationship.

A Bayesian approach was used to explore the genetic structure of our samples, using Bayesian Analysis of Population Structure (BAPS 3) software (<http://www.rni.helsinki.fi/~jic/bapspage.html>) (Corander et al., in press). BAPS 3, a program for Bayesian inference of the genetic structure in a population, treats both the allele frequencies of the molecular markers and the number of genetically diverged groups in a population as random variables. It uses stochastic optimization to infer the posterior mode of the genetic structure. Goodness-of-fit levels of the results are compared in terms of the "log ml-values" (natural logarithm of the marginal likelihood of the data) provided for any particular clustering solution. As recommended by the authors, the program was run with several values of K . The first runs were set with $K = 2$ to $K = 30$, K being the potential maximum number of genetically different groups (which will be called "clusters"), and was run five times for each value of K . For each K value (even the replicates of the same value) the program finds the optimal partitions with $k \leq K$. Then, because the best partitions seemed to lie somewhere below 30 we undertook many runs (at least 500) with $K = 30$. After having found the value of K which gave the highest "log ml" value, the F_{is} provided by this best partition was computed and its significance tested under FSTAT as above. The comparison between initial F_{is} (without partition) and the values obtained with the best partitions was also

undertaken with a Wilcoxon signed rank test for paired data, the pairing criterion being the individual loci (five data for each sample). BAPS exploration with $K = 50$ tested 1000 times was also applied on the whole data set (November 2000 and January 2001 samples pooled) in order to confirm the partitions found and to detect possible common clusters between the two samples. The last point was also checked with a tree construction approach on Cavalli-Sforza and Edwards (1967) chord distance between the clusters defined by BAPS. This distance is indeed the most appropriate for tree construction (Takezaki and Nei, 1996). The distances were computed by the GENETIX V. 4 software package (Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France). The distance matrix obtained was then used to build a dendrogram (neighbor joining method) (Saitou and Nei, 1987) using MEGA Version 3.1 (Kumar et al., 2004).

The homogeneity of age distribution among samples and clusters was tested with an analysis of variance with the software S-Plus 2000 professional release 3. For that, the factor "cluster" was nested into the factor "sample". The influence of fly's age on the infection with trypanosomes was investigated through a logistic regression with S-Plus 2000. The initial model was of the form:

$$\text{Trypanosome presence} \sim \text{Sample} + \text{Age} \\ + \text{Cluster (nested in sample)} + \text{Constant.}$$

A stepwise process allowed selecting for the minimum model and a Chi-square test was used to test the significance of the model (see S-PLUS, 2000 Guide to Statistics vol. 1 for more details). Comparisons of trypanosome prevalence between clusters in each of the two samples were done using Fisher's exact test (procedure STRUC in GENEPOP 3.3).

3. Results

3.1. Entomology

Of 7689 tsetse that were trapped (5664 females and 2025 males), all belonged to *G. p. palpalis*, based on morphological characters and on geographic distribution. Of these flies, 4421 were caught in November 2000 (sample 1, end of rainy season) and 3268 were trapped in January 2001 (sample 2, cold dry season). The apparent density averaged 3.45 tsetse/trap/day in the first survey, and 2.55 in the second, but varied between 1.4 and 8.5 according to the biotope. The sex ratio, 2.8 females to 1 male in the total sample, was typical of catches obtained with the Vavoua trap.

3.2. Epidemiological results

Among the 906 tsetse (799 females and 107 males) that were dissected, 25.06% were infected by trypanosomes. PCR techniques applied to these infected tsetse identified the following types of trypanosome infections: *Trypanosoma congolense* of the West African/Riverine forest type (*Tcf*, 6.5%), *Trypanosoma vivax* (*Tv*, 5.8%), *T. brucei* sensu lato (*Tb*,

Table 1
Characteristics of the five microsatellite loci used to study *G. p. palpalis* in Bonon, Côte d'Ivoire

Sample	Locus	Genic diversity	F_{is}	P -value	Frequency of most abundant allele	Number of alleles
1 (November 2000)	<i>Gpg55.3</i>	78.37	0.436	<0.001	0.44	17
	<i>Pgp24</i>	87.67	0.229	<0.001	0.26	20
	<i>Pgp13</i>	76.57	0.443	<0.001	0.42	16
	<i>Pgp11</i>	73.98	0.326	<0.001	0.4	10
	<i>Pgp1</i>	88.29	0.306	<0.001	0.23	14
2 (January 2001)	<i>Gpg55.3</i>	77.02	0.463	<0.001	0.44	17
	<i>Pgp24</i>	82.50	0.559	<0.001	0.30	13
	<i>Pgp13</i>	71.62	0.474	<0.001	0.44	13
	<i>Pgp11</i>	70.97	0.228	<0.001	0.47	11
	<i>Pgp1</i>	85.88	0.118	<0.001	0.29	16

2.6%), and *T. congolense* of the savannah type (*Tcs*, 0.70%). Among the tsetse infected with *T. brucei* s.l., only three individuals were infected with *T. b. gambiense* group 1, as indicated by the gambiense group 1 specific molecular marker. One of these flies had a salivary gland infection that was also detected by microscopy. Some of the trypanosomes observed by microscopy could not be identified by the molecular markers used (“undetermined” trypanosomes).

The mean age of the 770 females, determined by ovarian dissection, was 34.5 days. Age composition of these flies was 1.8% teneral (i.e. recently emerged but not fed), 14.2% nulliparous (no offspring deposited), 39% young females, and 45% old (i.e. more than 40 days).

3.3. Polymorphism, heterozygosity, null alleles, and short allele dominance at microsatellite loci

The number of alleles, at each locus, varied between 10 and 20 within each sample, and genic diversity varied between 70 and 88% (Table 1). Weir and Cockerham (1984) estimators indicated large heterozygote deficiencies within each sample (over all loci $F_{is} = 0.344$, $P < 0.001$, and $F_{is} = 0.366$, $P < 0.001$, for samples 1 and 2, respectively), with values of F_{is} by locus ranging from +0.12 to +0.56. All five loci deviated from Hardy–Weinberg expectations, showing a much lower number of heterozygotes than expected. No linkage disequilibrium was found between any pair of loci in either sample (data not shown). To explain these highly significant heterozygote deficiencies, both technical (e.g. undetectable nature of null alleles) and biological (e.g. Wahlund effect or inbreeding) factors must be taken into account.

If locus *Gpg55.3* is located on the X-chromosome (Solano et al., 2000; Gooding et al., 2004), and if null alleles were the cause of the heterozygote deficiencies, we would have expected to find null males. However, all males showed one band. In sample 2, under Brookfield hypothesis (Brookfield, 1996) with a frequency of null allele $r = 0.19$, 1.6 null males were expected (out of 8), whereas all males showed one band. The binomial P -value was 0.17 which was not significant, but was the minimum possible value given the small sample size. Even in females, the frequency of null alleles accounting for the observed deficiency would be so high (0.2) that null homozygotes (i.e. females showing no PCR product) would have been expected

(Brookfield, 1996), but none was observed. Thus, although the presence of null alleles could not be rejected, it is not likely that null alleles could explain the observed heterozygote deficiencies at this locus.

Primers were successfully redesigned for loci *Pgp24* and *Pgp13* (which showed the highest F_{is} values). However, we were not able to redesign efficient primers at the other loci. Using the new primers for locus *Pgp24* in sample 2, F_{is} decreased from +0.56 to +0.42, and at locus *Pgp13*, F_{is} decreased from +0.47 to +0.39. Despite this decrease in F_{is} , heterozygote deficiency remained significant at each locus (*Pgp24*, $F_{is} = 0.42$, $P < 0.0001$; *Pgp13*, $F_{is} = 0.39$, $P < 0.0001$).

There was a negative correlation between allele size and F_{is} value at four of the five loci (not at *Pgp11*), that is, high F_{is} values for short alleles, which can be called short allele dominance. The associated P -values appeared significant for both samples at locus *Pgp1*, and for either sample 1 or sample 2 at the three other loci (see Table 2). When the P -values were significant, they explained between 23 and 57% of the total variance of F_{is} at the locus.

We concluded that null alleles and/or short allele dominance may partly explain the overall heterozygote deficiency, but were not sufficient to explain it entirely, and that biological hypotheses had to be investigated.

Table 2

Results of regressions for each locus in each sample that showed a negative relationship between allele size and F_{is} value, as expected under the hypothesis of short allele dominance

Locus	Sample	Rho	P -value	R^2
<i>Gpg55.3</i>	1	−0.62	0.01	0.43
	1	−0.69	0.003	0.57
<i>Pgp13</i>	1	−0.48	0.06	0.19
	2	−0.61	0.03	0.35
<i>Pgp1</i>	1	−0.66	0.02	0.43
	2	−0.54	0.04	0.34

Numbers in bold represent significant P -values (<0.05) associated to the Spearman correlation coefficient (see Section 2 for details). R^2 is the coefficient of determination of the regression and represents the percentage of variance of F_{is} that is explained by the size of alleles, and hence illustrates with what magnitude short allele dominance explains the heterozygote deficits.

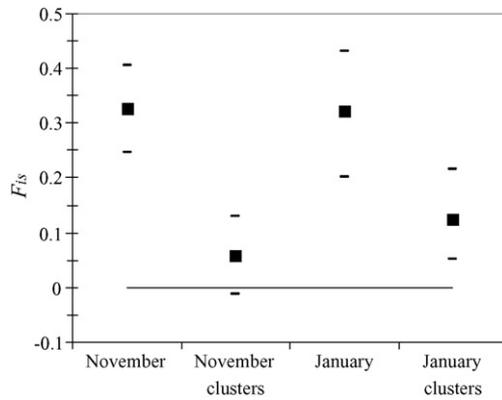


Fig. 2. F_{is} values estimated for both samples 1 and 2 *G. palpalis palpalis* before (November and January) and after BAPS partitioning of the data (clusters). 95% confidence intervals were obtained by bootstrapping over loci (See Goudet, 1995).

3.4. Geography

Our samples allowed us to test whether a Wahlund effect has occurred through artificially grouping populations from

different traps. To do this, tsetse were analysed according to the trap in which they were caught, considering each trap as a population. This was possible for 28 traps in which the number of females analysed was more than one. The overall F_{is} was 0.29 (0.22–0.38, 95% bootstrap confidence interval), showing no significant decrease compared to the pooled data ($F_{is} = 0.344$).

We also tried to group individual flies geographically, and we defined arbitrarily five zones: North, West, South, East, and the town of Bonon. Overall F_{is} was still 0.30 (0.24–0.38, 95% bootstrap confidence interval).

3.5. Identification of genetic groups within a sample

The best partitions obtained with the BAPS analysis provided 25 clusters in sample 1 and 21 clusters in sample 2. As shown in Fig. 2, the heterozygote deficiency dramatically dropped in both samples after clustering. In the partition obtained in sample 1, the F_{is} was not significantly different from zero (P -value = 0.07), but it remained significant in the partition obtained in sample 2 (P -value = 0.0001). In both

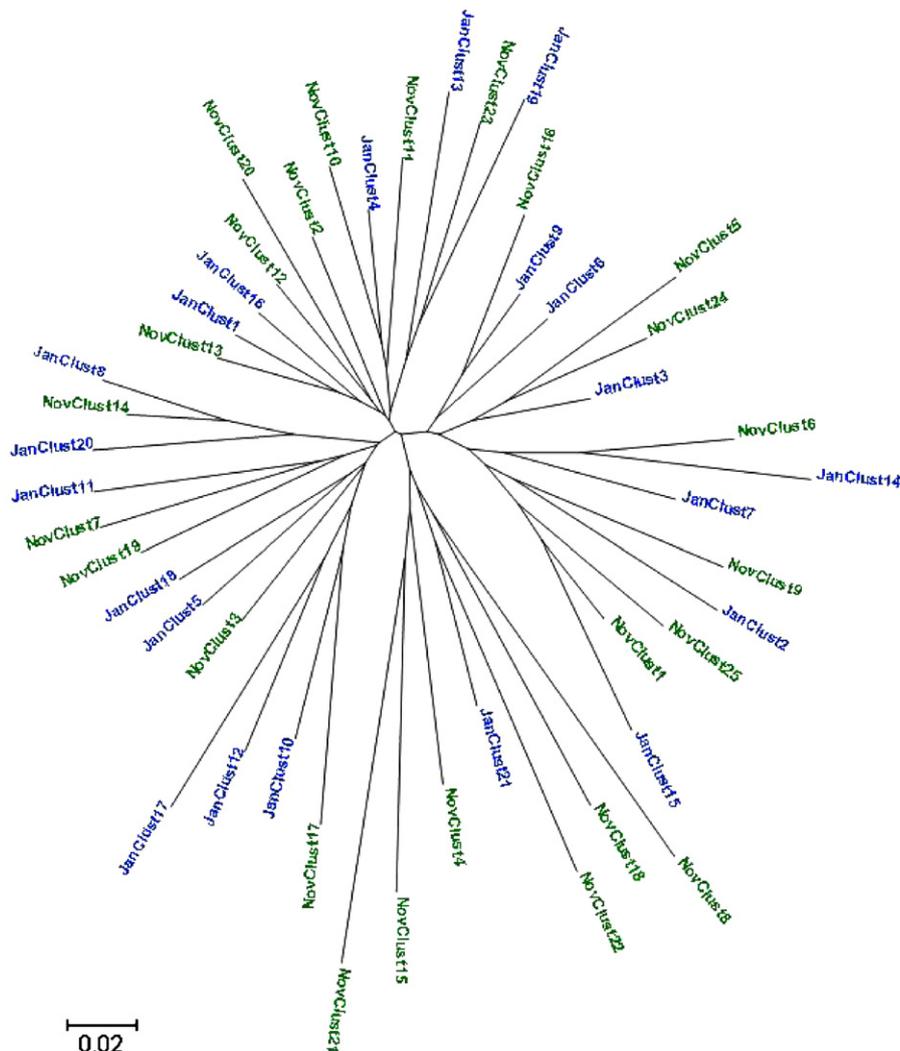


Fig. 3. Neighbour-joining tree obtained on Cavalli-Sforza and Edwards (1967) chord distances between BAPS clusters of *G. palpalis palpalis* from sample 1 (in green) and sample 2 (in blue).

samples the difference before and after clustering was significant (Wilcoxon signed rank test, P -value = 0.03).

The NJ tree obtained (Fig. 3) revealed no specific pattern, and except for rare pairs, no special relationship seemed to link samples 1 and 2 clusters. This was confirmed by the BAPS analysis undertaken on all individuals from both samples and pooled in a single one. The best partition obtained displayed 39 clusters. Less than one-third of these clusters (10) were heterogeneous with individuals from both samples, representing 46 individuals (among 225). This supports the lack of correspondence between clusters from samples 1 and 2. For both samples, the differentiation between clusters was $F_{st} = 0.29$ and 0.24 for samples 1 and 2, respectively, i.e. similar in magnitude.

3.6. Characteristics of the clusters

The composition of the 25 clusters of sample 1 and of the 21 clusters of sample 2 varied from 1 to 13 individuals per cluster, the mean being 5 for each of the two samples. Age was not significantly different between the two samples nor between clusters within samples (ANOVA, P -value > 0.16). The logistic regressions did not highlight any specific pattern except that more trypanosomes were found in sample 2 (prevalence = 0.41) than in sample 1 (prevalence = 0.29) (P -value = 0.023). Nevertheless, sample sizes were small and might have lowered the power of detection of some other patterns. Ignoring age, Fisher exact tests revealed that trypanosome prevalence (by microscopic examination) was heterogeneous among clusters from sample 1 (P -value = 0.001), but not from sample 2 (P -value = 0.169). Indeed some clusters had no trypanosome infections, or very low infection rates (e.g. cluster 3 from sample 1 and cluster 13 from sample 2), whereas others showed very high infection rates (above 40%, e.g. clusters 7 and 15 of sample 1 containing 13 individuals each, and clusters 3, 6, and 8 from sample 2 containing 7, 9, and 10 individuals each, respectively). Some of the trypanosomes found by microscopy could not be identified and were reported as “undetermined”. With molecular determination of trypanosome species, because of small sample sizes, statistical testing was rarely possible. Some clusters harboured only one trypanosome type as recognised by PCR (e.g. clusters 8 and 9 from sample 1 and cluster 7 from sample 2), whereas others were shown to be infected by several trypanosome species or groups. For *T. brucei* s.l. the heterogeneity of prevalence was marginally significant in sample 2 (P -value = 0.042). For *T. congolense* of the Riverine Forest type, the distribution seemed homogeneous among clusters (P -value > 0.7 in both samples). For *T. vivax*, prevalence was heterogeneous only in sample 1 (P -value = 0.036) but not in sample 2 (P -value = 0.93). For the undetermined trypanosomes, no differences were observed in any sample (P -value > 0.6). Bloodmeal sources of some tsetse could be identified as coming from animals, e.g. clusters 8, 12, and 25 from sample 1, and cluster 3 from sample 2. All those clusters having such identified bloodmeal sources showed trypanosome infections.

4. Discussion

The present study was undertaken to gain knowledge of the population structure of *G. p. palpalis*, the main tsetse species found at Bonon, a recently described, very active HAT focus in west-central Côte d’Ivoire. The rationale for this work is the assumption that a knowledge of the genetic structure of a vector population will help provide a sound basis for understanding an epidemic and will contribute to a rational control operation, especially one that involves a genetic approach to vector control. Allele frequencies were determined at five microsatellite loci in two temporal samples. Highly significant heterozygote deficits, i.e. strongly positive values of F_{is} , the within-population index measuring deviation from random mating, were observed at the five microsatellite loci in each of the two samples taken three months apart.

These high F_{is} values were not expected and are not usual values found in other Dipteran groups. However, molecular data on tsetse have been very scarce up to now, and whenever they exist they are mostly directed towards *morsitans* group species. For instance, in the review of Gooding and Krafsur (2005), there is basically random mating within populations but not among (in *G. pallidipes* and *G. morsitans*), which was found surprising by the authors compared to the reported vagility of tsetse. The first preliminary study on natural populations of *G. palpalis gambiensis* in Burkina Faso using microsatellite markers showed high F_{is} values which were explained by the existence of a Wahlund effect (Solano et al., 2000). These authors did not find so high F_{is} values in other localities, for instance in Senegal and in another area of Burkina Faso (Solano et al., 1999), where there was no indication of deviation from within population random mating. In one of the few studies on tsetse using microsatellite DNA loci (Luna et al., 2001) where they developed 13 microsatellite loci and tested them on several lab reared tsetse taxa, the only field population of *G. palpalis gambiensis* they tested (coming from Burkina Faso, 30 individuals) showed heterozygote deficiency at 11 out of 12 loci, with 3 loci showing significant departure from H–W expectations. Elsen et al. (1994), using multilocus enzyme electrophoresis on laboratory-reared *G. palpalis gambiensis*, also found departure from Hardy–Weinberg expectations due to significant heterozygote deficiencies at some of the loci studied.

The situation here may also be complicated by the fact that the study area is located close to the northern limit of *G. p. palpalis*, hence to the limit between the two subspecies of *G. palpalis* (Challier et al., 1983). The possibility cannot be ruled out that some of the clusters contain individuals coming from crosses of the two subspecies, although in the laboratory hybridisation between these subspecies leads to sterile males and fertile females (Southern, 1980; Gooding, 1997), and thus females provide a bridge for introgression of genes. The possibility that *G. palpalis* indeed represents a species complex should be examined on a wider geographic scale.

In the present study, null alleles were demonstrated at two loci by using redesigned primers, but their frequency was too low to fully explain the heterozygote deficits. Short allele

dominance was also observed and accounted probably for part of the total F_{is} . Geographic structuring did not explain the observed deficiencies, regardless of the geographic scale used in the analysis. Even at the scale of a single trap, for some of the traps, heterozygote deficiencies remained very high and statistically significant.

The significant decrease in heterozygote deficits, leading eventually to a non-significant value in sample 1, and observed also in at least 39 of the clusters that were defined by the BAPS approach, strongly suggests that the Vavoua traps attracted tsetse flies from numerous, strongly differentiated units (i.e. sub-populations or groups). The initially high heterozygote deficits are thus most likely explained by a strong Wahlund effect and, to a lesser extent, by the presence of null alleles and short allele dominance. These tsetse groups would be of small size, which is generally acknowledged as being typical for tsetse (e.g. Gooding and Krafsur, 2005), and would favour genetic differentiation due to genetic drift. Tsetse flies may be vagile enough (see Leak, 1999 for review) to enable the trapping system to sample many of these genetically differentiated groups, leading repeatedly to high F_{is} values. Genetic drift is not necessarily contradicted by the reported vagility of *G. palpalis* since tsetse may not use their maximal capacities of dispersal if they do not need to: very recent data on mark-release-recaptures of *G. palpalis* in Burkina Faso showed that when tsetse are released in very favorable area (i.e. an area with humidity and shade typical of *G. palpalis*), a statistically significant number stays within this area, whereas when they are released in a degraded area, they tend to look for a favorable area and then diffuse on much longer distances (J. Bouyer, personal communication). The Wahlund effect found here then reveals the artificial grouping in the traps of several different genetic groups. In addition, most of the traps being placed only on sites daily frequented by HAT patients, they are likely to catch mostly tsetse that are looking for a bloodmeal. If there is sub-structuring, tsetse from different subgroups would be artificially mixed in these traps when they come to hunt. If traps were placed near larviposition sites, in more conserved areas, F_{is} values might be lower, and that is a possibility that will deserve further work. Depending on availability of hosts in the vicinity of these different subgroups, the different clusters would appear to be differentially infected, as may be the case in this study. It is not known whether these genetically different populations differ in their vectorial competence and/or host preferences, as previously suggested with other field-caught tsetse (Solano et al., 2000) and in laboratory experiments (Elsen et al., 1994).

The fact that little correspondence was found between the clusters defined in the two samples suggests that, because of the number of groups and the number of traps from which the samples originated (57 for sample 1, 47 from sample 2, 14 of which were exactly at the same place), the probability of catching individuals from the same sub-population during two temporally distinct sampling periods was very low. Possible exceptions include cases where traps were placed at precisely the same locations in both sampling periods and attracted flies from habitats that supported the same sub-

populations, despite any microhabitat changes that may have occurred between sampling dates. It has also to be said that the number of clusters found here does not necessarily represent the true number of existing sub-populations, because this will depend very much on the spatial distribution of the traps within the area together with the dispersal capacity of the tsetse, and also because the number of BAPS clusters might have considerably been increased by short allele dominance at several loci.

Our suggestion that the *G. p. palpalis* population at Bonon composed of small units is consistent with the reported vulnerability of tsetse to trapping, which fortunately, is an efficient control method (Laveissière and Penchenier, 2005; Kuzoe and Schofield, 2005). The suggested population structure may also account for both the reported ability of tsetse populations to recover rapidly from low population densities and for their ability to rapidly re-invade controlled areas (Torr et al., 2005). This result also suggests that traps may result in a sampling bias, and that other sampling methods (e.g. Muzari and Hargrove, 2005) should be considered in order to look for the real reproductive units. Finally, the genetic pattern observed in *G. palpalis* populations may be explained either by very small sub-population sizes and/or by a strong phylopatric behaviour, as larval pheromones have been suspected to occur (Nash et al., 1976; Gouteux et al., 1983).

Since the HAT focus of Bonon has been only recently described, and since HAT displays an epidemic pattern at Bonon (Djè et al., 2002), an understanding of the dynamics and composition of the local tsetse populations is undoubtedly a crucial issue. Such knowledge may elucidate the importance of HAT, which seems to be directly linked to human and environmental changes in this area (Courtin et al., 2005). Studies on both subspecies of *G. palpalis* s.l. (Solano et al., 2000 and the present one) concluded that the presence of a Wahlund effect explains, at least in part, the genetic structure of natural populations of this species. Future work ought to design other sampling strategies to access the real reproductive units (sub-populations) of this important vector species. Furthermore, future work needs to examine additional molecular and morphometric markers on the same individual to shed light on this issue. Finally there is a need to extend these studies to other tsetse species in other geographic locations to help determine on where and how to eliminate tsetse.

Acknowledgments

The work was funded by IRD, with additional support by FAO/IAEA Co-ordinated Research Program no. 11421 RBF, French Service de Coopération et d'Action Culturelle Abidjan, and the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR). We also wish to thank the "tsetse and trypanosomosis" Research Unit and the Direction of Institut Pierre Richet, and Dr. Jeremy Bouyer (CIRAD-EMVT/CIRDES) and two anonymous referees for helpful comments on the manuscript.

References

- Awono-Ambene, H.P., Kengne, P., Simard, F., Antonio-Nkondjio, C., Fontenille, D., 2004. Description and bionomics of *Anopheles (Cellia) ovengensis* (Diptera: Culicidae), a new malaria vector species of the *Anopheles nili* group from south Cameroon. *J. Med. Entomol.* 41, 561–568.
- Brookfield, J.F.Y., 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol. Ecol.* 5, 453–455.
- Cavalli-Sforza, L.L., Edwards, A.W.F., 1967. Phylogenetic analysis: model and estimation procedures. *Am. J. Hum. Genet.* 19, 233–257.
- Challier, A., 1965. Amélioration de la méthode de détermination de l'âge physiologique des glossines. *Bull. Soc. Pathol. Exot.* 58, 250–259.
- Challier, A., Gouteux, J.P., Coosemans, M., 1983. La limite géographique entre les sous-espèces *Glossina palpalis palpalis* (Rob.-Desv.) et *Glossina palpalis gambiensis* Vanderplanck en Afrique occidentale. *Cah. O.R.S.T.O.M., Sér. Entomol. Méd. Parasitol.* 21, 207–220.
- Corander, J., Marttinen, P., Mäntyniemi, S. Bayesian identification of stock mixtures from molecular marker data. *Fish Bull.*, in press.
- Courtin, F., Jamonneau, V., Oké, E., Oswald, Y., Coulibaly, B., Dupont, S., Doumenge, J.P., Cuny, G., Solano, P., 2005. Trying to understand presence/absence of human African Trypanosomiasis in Côte d'Ivoire: analysis of the actors of the pathogenic system. *Int. J. Health Geogr.* 4, 27.
- Cuisance, D., Fevrier, J., Dejardin, J., Filledier, J., 1985. Dispersion linéaire de *Glossina palpalis gambiensis* et *G. tachinoides* dans une galerie forestière en zone soudano-guinéenne. *Rev. Elev. Méd. Vét. Pays Trop.* 38, 153–172.
- De Meeüs, T., Humair, P.F., Grunau, C., Delaye, C., Renaud, F., 2002. Non mendelian transmission of alleles at microsatellite loci: an example in *Ixodes ricinus*, the vector of Lyme disease. *Int. J. Parasitol.* 34, 943–950.
- Djè, N.N., Miézan, T.W., N'Guessan, P., Brika, P., Doua, F., Boa, F., 2002. Distribution géographique des trypanosomés pris en charge en Côte d'Ivoire de 1993 à 2000. *Bull. Soc. Pathol. Exot.* 95, 359–361.
- Elsen, P., Roelants, P., De Lil, E., Dujardin, J.P., Le Ray, D., Claes, Y., 1994. Cytogenetic and isozymic comparisons of two laboratory lines of *Glossina palpalis gambiensis*. *Ann. Trop. Med. Parasitol.* 88, 511–522.
- Gooding, R.H., 1996. Genetic variation in arthropod vectors of disease-causing organisms: obstacles and opportunities. *Clin. Microbiol. Rev.* 9, 301–320.
- Gooding, R.H., 1997. Genetic analysis of hybrid sterility in crosses of the tsetse flies *Glossina palpalis palpalis* and *Glossina palpalis gambiensis* (Diptera: Glossinidae). *Can. J. Zool.* 75, 1109–1117.
- Gooding, R.H., Krafusur, E.S., 2005. Tsetse genetics: contributions to biology, systematics, and control of tsetse flies. *Ann. Rev. Entomol.* 50, 101–123.
- Gooding, R.H., Solano, P., Ravel, S., 2004. X-chromosome mapping experiments suggest occurrence of cryptic species in the tsetse fly *Glossina palpalis palpalis*. *Can. J. Zool.* 82, 1902–1909.
- Goudet, J., 1995. FSTAT (Version 1.2): a computer program to calculate *F*-statistics. *J. Hered.* 86, 485–486.
- Gouteux, J.P., Laveissière, C., Boreham, P.F.L., 1982. Ecologie des glossines en secteur pré-forestier de Côte d'Ivoire. 2. Préférences trophiques de *G. palpalis* s.l. *Cah. O.R.S.T.O.M., Sér. Entomol. Méd. Parasitol.* 20, 3–18.
- Gouteux, J.P., Laveissière, C., Couret, D., 1983. Ecologie des glossines en secteur pré-forestier de Côte d'Ivoire. 5. Les lieux de reproduction. *Cah. O.R.S.T.O.M., Sér. Entomol. Méd. Parasitol.* 21, 3–12.
- Jamonneau, V., Ravel, S., Koffi, M., Zeze, D.G., Kaba, D., N'Dri, L., Coulibaly, B., Cuny, G., Solano, P., 2004. Mixed trypanosome infections in tsetse and pigs and their epidemiological significance in a sleeping sickness focus in Côte d'Ivoire. *Parasitology* 129, 693–702.
- Jannin, J.G., 2005. Sleeping sickness—a growing problem? *Br. Med. J.* 331, 1242.
- Kaba, D., Djè, N.N., Courtin, F., Oke, E., Koffi, M., Garcia, A., Jamonneau, V., Solano, P., 2006. Impact de la guerre sur l'évolution de la THA dans le Centre-Ouest de la Côte d'Ivoire. *Trop. Med. Int. Health* 11, 136–143.
- Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* 5, 150–163.
- Kuzoe, F.A.S., Schofield, C.J., 2005. Strategic review of traps and targets for tsetse and African Trypanosomiasis control. *TDR/IDE/TRY/05.1*, 58 pp.
- Lanzaro, G.C., Warburg, A., 1995. Genetic variability in phlebotomine sandflies: possible implications for Leishmaniasis epidemiology. *Trends in Parasitology* 11, 151–154.
- Laveissière, C., Grébaud, P., 1990. Recherches sur les pièges à glossines. Mise au point d'un modèle économique: le piège "Vavoua". *Trop. Med. Parasitol.* 41, 185–192.
- Laveissière, C., Penchenier, L., 2005. Manuel de lutte contre la maladie du sommeil. IRD ed. Institut de Recherche pour le Développement, Collections Didactiques, Paris, 366 pp.
- Leak, S.G.A., 1999. Tsetse Biology and Ecology: Their Role in the Epidemiology and Control of Trypanosomiasis. CABI Publishing, Wallingford.
- Louis, F.J., 2001. Les raisons techniques de la ré-émergence de la maladie du sommeil. *Méd. Trop.* 61, 425–431.
- Luna, C., Bonizzoni, M., Cheng, Q., Aksoy, S., Zheng, L., 2001. Microsatellite polymorphism in the tsetse flies (Diptera: Glossinidae). *J. Med. Entomol.* 38, 376–381.
- Majiwa, P.A.O., Thatthi, R., Moloo, S.K., Nyeko, J.H.P., Otieno, L.H., Maloo, S., 1994. Detection of trypanosome infections in the saliva of tsetse flies and buffy-coat samples from antigenemic but aparasitaemic cattle. *Parasitology* 108, 313–322.
- Masiga, D.K., Smyth, A.J., Hayes, P., Bromidge, T.J., Gibson, W.C., 1992. Sensitive detection of trypanosomes in tsetse flies by DNA amplification. *Int. J. Parasitol.* 22, 909–918.
- Muzari, M.O., Hargrove, J.W., 2005. Artificial larviposition sites for field collections of the puparia of tsetse flies *Glossina pallidipes* and *G. m. morsitans* (Diptera: Glossinidae). *Bull. Entomol. Res.* 66, 221–229.
- Nash, T.A.M., Treweek, M.A., Moloo, S.K., 1976. Observations on the free larval stage of *G. morsitans morsitans* Westw. (Diptera: Glossinidae): the possibility of a larval pheromone. *Bull. Entomol. Res.* 66, 17–24.
- Raymond, M., Rousset, F., 1995. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* 86, 248–249.
- Reid, R.S., Kruska, R.L., Deichmann, U., Thornton, P.K., Leak, S.G.A., 2000. Human population growth and the extinction of the tsetse fly. *Agric. Ecosyst. Env.* 77, 227–236.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- S-PLUS, 2000. Guide to Statistics, vol. 1. Data Analysis Products Division, MathSoft, Seattle, WA.
- Solano, P., Argiro, L., Reifenberg, J.M., Yao, Y., Duvallet, G., 1995. Field application of the polymerase chain reaction (PCR) to the detection and characterization of trypanosomes in *Glossina longipalpis* in Côte d'Ivoire. *Mol. Ecol.* 4, 781–785.
- 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., de La Rocque, S., de Meeüs, T., Cuny, G., Duvallet, G., Cuisance, D., 2000. Microsatellite DNA markers reveal genetic differentiation among populations of *Glossina palpalis gambiensis* collected in the agro-pastoral zone of Sideradougou, Burkina Faso. *Insect Mol. Biol.* 9, 433–439.
- Solano, P., de La Rocque, S., Cuisance, D., Geoffroy, B., de Meeus, T., Cuny, G., Duvallet, G., 1999. Intraspecific variability in natural populations of *Glossina palpalis gambiensis* from West Africa, revealed by genetic and morphometric analyses. *Med. Vet. Entomol.* 13, 401–407.
- Solano, P., Koné, A., Garcia, A., Sané, B., Michel, V., Michel, J.F., Coulibaly, B., Jamonneau, V., Kaba, D., Dupont, S., Fournet, F., 2003. Rôle des déplacements des malades dans l'épidémiologie de la Trypanosomose Humaine Africaine dans le foyer de Bonon, Côte d'Ivoire. *Méd. Trop.* 63, 577–582.
- Southern, D.I., 1980. Chromosome diversity in tsetse flies. In: Blackman, R.L., Hewitt, G.M., Ashburner, M. (Eds.), *Insect Cytogenetics, Symposia of the Royal Entomological Society of London*, No. 10. Blackwell Scientific Publications, Oxford, U.K., pp. 225–243.
- Takezaki, N., Nei, M., 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144, 389–399.
- Torr, S.J., Hargrove, J.W., Vale, G.A., 2005. Towards a rational policy for dealing with tsetse. *Trends Parasitol.* 21, 537–541.
- Truc, P., Ravel, S., Jamonneau, V., N'Guessan, P., Cuny, G., 2002. Genetic variability within *Trypanosoma brucei gambiense*: evidence for the

- circulation of different genotypes in human African trypanosomiasis patients in Côte d'Ivoire. *Trans. R. Soc. Trop. Med. Hyg.* 96, 1–4.
- Walsh, P.S., Metzger, D.A., Higuchi, R., 1991. Chelex[®] 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10, 506–513.
- Wattier, R., Engel, C.R., Saumitou-Laprade, P., Valero, M., 1998. Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Mol. Ecol.* 7, 1569–1573.
- Weir, B.S., Cockerham, C.C., 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- WHO, 2001. African Trypanosomiasis or sleeping sickness. World Health Organisation fact sheet 259, <http://www.who.int/mediacentre/factsheets/fs259/en>.