Trypanosoma brucei from pigs in sleeping sickness foci from Côte d'Ivoire is structured into clonal and strongly subdivided populations

Martial Kassi N'Djetchi [a, 1, *], Thierry De Meeûs [b, 1], Barkissa Mélika Traoré [a, 1], Sophie Ravel [b], Adeline Ségard [b], Jacques Kaboré [d], Abe Allépo Innocent [a], Dramane Kaba [e], Djakaridja Berté [e], Thomas Konan [a], Bamoro Coulibaly [e], Pascal Grébaut [b], Geraldine Bossard [b,c], Bruno Bucheton [b], Mathurin Koffi [a, 2], Vincent Jamonneau [b, e, 2]

[1]Joint first authors

[2]Joint last authors

[a]Unité de Recherche en Génétique et Epidémiologie moléculaire, Laboratoire de biodiversité et Gestion durable des écosystèmes tropicaux, Université Jean Lorougnon Guédé, BP 150 Daloa, Côte d'Ivoire.

[b]Intertryp, IRD, CIRAD, University of Montpellier, France.

[C] CIRAD, UMR INTERTRYP, F-34398, Montpellier, France

[d]Unité de recherches sur les maladies à vecteurs et biodiversité, Centre International de Recherche-Développement sur l'Elevage en zones Subhumides, 01 BP 454 Bobo-Dioulasso 01, Burkina Faso.

[e]Unité de Recherche Trypanosomoses, Institut Pierre Richet, 01 BP 1500, Bouaké, Côte d'Ivoire.

* Corresponding author: ndjetchi2015@gmail.com

Abstract

Human African trypanosomiasis (HAT), or sleeping sickness, is currently targeted for elimination. The etiologic agent of HAT is a trypanosome belonging to the species *Trypanosoma brucei* (*Tb*) s.l., a unicellular parasite transmitted by tsetse flies. *Tb* s.l. consists of three subspecies: *T. b. brucei* (*Tbb*), *T. b. gambiense* (*Tbg*) and *T. b. rhodesiense* (*Tbr*). These subspecies are morphologically indistinguishable and classified according to host in which they are found, type of disease and geographical distribution. During the last few decades, there has been considerable effort to genetically characterize *Tb* s.l. isolated from domestic and wild animals in order to better evaluate the impact of animal reservoirs on the epidemiology of HAT. To assess genetic diversity of *Tb* s.l. strains circulating in three endemic or historical HAT foci in Côte d'Ivoire, we conducted a population genetics study of these parasites. Biological and isolated stock samples collected from pigs and reference stocks were tested with the primers of the *Trypanosoma gambiense*-specific-glycoprotein gene (TgsGP) and were genetically characterized with eighteen microsatellite primers. TgsGP positive samples did not fit into *Tbg* as regard to their microsatellite profile. We also found that in Ivoirian foci, *Tbb* populations (animal trypanosomes) were structured as several strongly isolated units that propagate clonally. This is in variance with other published data on that subspecies. This study confirms the need to develop better tools to explore the relationships between *Tbb* and *Tbg* and to study the epidemiological role of potential animal reservoir for *Tbg*.

Keywords: Human African trypanosomiasis; Microsatellite markers; Animal reservoir; Population genetics; Dispersal, Côte d'Ivoire.

Introduction

Human African trypanosomiasis (HAT), also known as sleeping sickness, is currently targeted for elimination (Franco et al., 2024). Its potential for devastating epidemics was demonstrated throughout the 20th century (WHO, 2013). The causative agent is *Trypanosoma brucei* (*Tb*) s.l., a unicellular parasite transmitted by tsetse flies of the genus *Glossina*. *Trypanosoma brucei* infects human as well as a variety of domestic and wild animals in sub-Saharan Africa (Lejon et al., 2025).

Trypanosoma brucei s.l. was subdivided into three subspecies based on extrinsic criteria (especially host range, pathogenicity in human and geographical distribution): *T. b. gambiense* (*Tbg*) and *T. b. rhodesiense* (*Tbr*) are both human-infective and responsible for a chronic form in West and Central Africa and an acute form in East Africa, respectively (Lejon et al., 2025). *Trypanosoma brucei brucei* (*Tbb*), theoretically non-pathogenic to humans, is a parasite of domestic and wild animals causing animal African trypanosomosis (AAT or nagana) throughout the tsetse region of Africa (Hoare, 1972). With the advent of molecular methods, *Tbg* was subdivided into *Tbg* group 1 (*Tbg*1), a homogenous genetic entity responsible of most HAT cases (Gibson et al., 1986) and the rarest *Tbg* group 2 (*Tbg*2) recently defined as all human-infective *Tb* trypanosomes from West and Central Africa that do not fit into *Tbg*1 using molecular markers (Jamonneau et al., 2019).

With 747 and 799 cases reported in 2021 and 2022 respectively, *gambiense* HAT (g-HAT) responsible for 94% of the total number of reported cases, is currently targeted for interruption of transmission (Franco et al., 2024). However, the epidemiological role of domestic or wild animal reservoir which could compromise this objective, is still under debate (Koffi et al. 2015; Büscher et al., 2018; Mehlitz et al., 2019).

In Côte d'Ivoire, important control efforts have led to a significant reduction of the prevalence of the disease (Dje et al., 2002; Kaba et al., 2006; Courtin et al., 2010; Koffi et

al., 2016; Koné et al., 2020). Thanks to these efforts, the country has achieved the elimination of g-HAT as a public health problem in 2020 (Kaba et al., 2023), although a few cases were still diagnosed these last years (Kone et al., 2021). Potential animal reservoirs of *Tbg* that have been described for a long time in HAT foci (Mehlitz, 1986; Jamonneau et al., 2004) are now suspected to be partly responsible for the persistence of the disease (Kambire et al., 2012; Koffi et al., 2016; Kaba et al., 2023). In recent studies conducted in the Bonon and Sinfra endemic foci and the Vavoua historical focus, high prevalence of *Tb* s.l. was observed in domestic animals, mainly in pigs, due to a higher exposure of these animals to tsetse flies (N'Djetchi et al., 2017; Traoré et al., 2021). However, discordant results were obtained with the molecular and serological tools used. The specificity of the PCR targeting the *Trypanosoma gambiense*-specific-glycoprotein gene (TgsGP) described as a suitable method for the specific detection of *Tbg* (Gibson et al. 2010; Radwanska et al. 2002b) was even questioned and both the existence and the epidemiological role of an animal reservoir for *Tbg* still remains unclear (N'Djetchi et al., 2017; Traoré et al., 2021).

The aim of the present study was to explore the population genetics of trypanosomes collected in the field from pigs from these three Ivorian gHAT foci (Bonon, Sinfra and Vavoua), and compare it to several reference stocks of *Tb* s.l. Our findings will bring clarifications on the relationships between *Tbb* and *Tbg*, their reproductive mode, dispersal and on their epidemiology.

Material and Methods

Ethical statement

Sample collection was conducted by a veterinary of the Laboratoire National d'Appui au Développement Rural (Ministry of Agriculture) within the framework of epidemiological surveillance activities supervised by the HAT National Elimination Program (HAT NEP).

Local authorities require no ethical statement. Any veterinarian may carry out blood sampling on domestic animals, with the authorization of the owner. No samples other than those for routine screening and diagnostic procedures were collected. Breeders gave their consent for animal sampling after explaining the objectives of the study. A deworming treatment (Bolumisol, Laprovet) was freely provided to all pigs sampled.

Study areas

All samples included in this study were from Bonon and Sinfra endemic HAT foci and from the Vavoua historical HAT focus in the Western Center part of Côte d'Ivoire (Figure 1). These foci fall into the distribution area of tsetse (*Glossina*) (Laveissière et al., 1981; Ravel et al., 2007; Kaba et al., 2021). In these zones, AAT is mostly due to *Tbb* and *T. congolense* that mainly affect both pigs and cattle (N'Djetchi et al., 2017; Traoré et al., 2021).

Control efforts, conducted from 1992 until present, largely contained the HAT epidemic in Sinfra (1992-1997) and in Bonon (1998-2004) (Dje et al., 2002; Kaba et al., 2006; Kambire et al., 2012; Kaba et al, 2021). Nevertheless, few cases are still diagnosed each year (Kambire et al., 2012; Koffi et al., 2016; Koné et al., 2021; Kaba et al., 2023). The Vavoua HAT focus was epidemic in the 1970's (Duvallet & Stanghellini, 1979) and thanks to control efforts, less than five HAT cases were reported yearly at the end of the 1990s (Dje et al., 2002). From the early 2000s, only four cases were still diagnosed (Simarro et al., 2010) and the last one was reported in 2011 (Traoré et al., 2021).

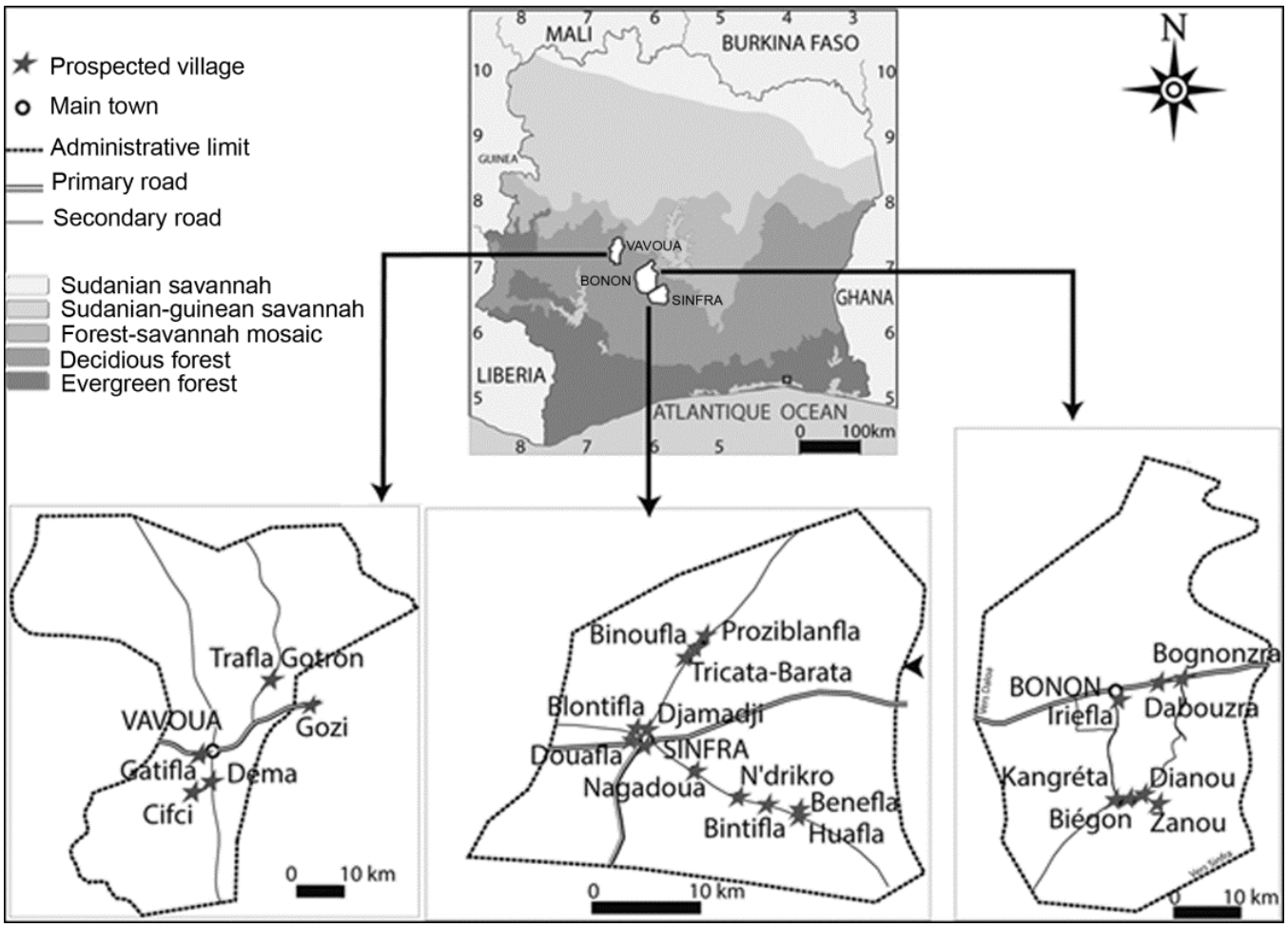

Figure 1: Study areas and sites of sampling of *Tb* s.l. from pigs. This figure was created by the mapping service of our team based at Institut Pierre Richet (Bouake, Côte d'Ivoire).

Sampling

The number and the description of the selected field samples according to the focus, the year, the sampling method and the village are given in Table 1. Regarding the Vavoua focus (2017), we used 20 biological samples (BS) consisting of 500 µL of buffy coat obtained from 9 mL of heparinised blood and 20 isolated stock samples (IS) already described in Traoré et al. (2021). Both selected BS and IS were positive with the TBR1-2 PCR (Moser et al., 1989) specific of *Tb* s.l.. All BS were negative for the single round PCR targeting the TgsGP gene specific of *Tbg* (Radwanska et al., 2002). Five of the IS were TgsGP PCR

positive with one giving a high intensity band and four a low intensity band (Traoré et al., 2021).

In the Bonon (2015), Sinfra (2016) and in the Sinfra foci (2017), the sample selection was also performed on BS (11 and 29, respectively) and IS (3 and 17, respectively) collected from pigs in the frame of studies on AAT epidemiology as previously described (Traoré et al., 2021). All these pigs tested positive for the trypanosome-detection buffy coat technique (Murray et al., 1997) and both BS and IS were positives to the TBR1-2 PCR but negative to the single round TgsGP PCR.

Considering that population differentiation becomes significant between subsamples from the same site but separated by six months (Kaboré et al., 2011), we split subsamples separated by more than six months as belonging to different cohorts (C). Sampling extended from December 2015 to March 2017. This resulted in two different cohorts for the present study: C1 (December 2015 to April 2016) and C3 (January to March 2017). For C1, sampling corresponded to Bonon in December 2015 and Sinfra in April 2016. For C3, sampling corresponded to Sinfra in January 2017, and to Vavoua in February and March 2017 (Table 1).

DNA from 27 reference isolated stocks of *Tbb* (n=10), *Tbg1* (n=13) and *Tbg2* (n=4) were also included in this study (Table 2).

Table 1**:** Description and number of samples used for the amplification of microsatellite loci of *Trypanosoma brucei sp* isolates. Data are given regarding focus, year of sampling and cohort. We considered two different periods of six months (Kaboré et al.,2011), the first corresponding to samples collected between December 2015 and May 2016 (cohort: C1) and the second corresponding to samples collected between December 2016 and May 2017 (Cohort: C3). The number of samples regarding the village, the sampling method (BS: biological samples, IS: isolated stock samples), as well as the number of samples that had at least 70% of interpretable loci used for the population genetics parameter estimates (N70%) are also indicated.

Focus	**Year**	**Cohort**	**Village**	**BS**	**IS**	**N70%**
			Biegon	1	1	1
			Dabouzra	0	2	1
Bonon	2015	1	Dianou	1	0	1
			Iriefla	2	0	2
			Kangreta	5	0	3
			Zanou	1	0	0
			Bognonzra	1	0	0
			Total	11	3	
			Douafla	2	0	1
			Binoufla	0	1	1
Sinfra	2016	1	Nagadoua	5	0	3
			N'Drikro	3	0	1
			Prozibla	1	0	1
			Tracata-B	1	0	1
			Total	12	1	
			Benefla	1	1	1
			Bentifla	2	3	5
			Blontifla	1	1	2
			Djamadji	2	1	1
Sinfra	2017	3	Douafla	2	0	2
			Huafla	2	3	3
			Nagadoua	0	4	3
			N'Drikro	4	2	3
			Sinfra	3	1	2
			Total	17	16	
			Dema	3	1	2
			Gatifla	6	2	5
Vavoua	2017	3	Gozi	6	8	11
			PK14	1	4	4
			Trafla	2	5	6
			CIFCI	2	0	0
			Total	20	20	
Total				**60**	**40**	

Table 2: Reference stocks of *Trypanosoma brucei sp* used for microsatellite loci characterisation, and taxonomic identification.

Stocks	Host	Origine	Year	Species	References
B4F303	Human	CI	2004	*Tbg*1	Koffi et al. (2007)
107_4	Human	CI	1997		Jamonneau et al. (2000)
B4U163	Human	CI	2004		Koffi et al. (2007)
JUA	Human	Cam	1979		Stevens et al. (1992)
A005	Human	Cam	1988		Stevens et al. (1992)
OK/ITMAP	Human	Con	1974		Stevens et al. (1992)
P26F	Pig	Cam	1998		Nkinin et al. (2002)
D12k	Sheep	Con	1980		Godfrey et al. (1990)
DIBBYEUG	Human	CI	1998		Jamonneau et al. (2000)
B4_K5	Human	CI	2004		Koffi et al. (2009)
B4_Y221	Human	CI	2004		Koffi et al. (2009)
B4C191	Human	CI	2004		Kaboré et al. (2018)
T59_3	Human	CI	2002		Koffi et al. (2009)
ABBA	Human	CI	1983	*Tbg*2	Hide et al. (1990)
MURAZ03	Human	CI	1979		Godfrey et al. (1987)
TH2	Human	CI	1978		Mehlitz et al. (1982)
FEOS	Human	Togo	1961		Tait et al. (1984)
SW94_87	Pig	DRC	1987	*Tbb*	Truc et al. (1991)
DC8	Pig	CI	2001		Jamonneau et al. (2004)
DC9	Pig	CI	2001		Jamonneau et al. (2004)
EATRO1125	Antilope	Ug	1966		Gibson et al. (1980)
P11F	Pig	Cam	1998		Simo et al. (2005)
P17F	Pig	Cam	1998		Nkinin et al. (2002)
GAOUA89	Zebu	BF	1989		Simo et al. (2005)
STIB215	Lion	Tan	1971		Gibson et al. (1980)
TSW65	Pig	CI	1982		Stevens et al. (1992)
TSW53	Pig	CI	1982		Stevens et al. (1992)

Microsatellite genotyping

DNA from BS was extracted from 500 µL of buffy coat using the DNeasy Blood and Tissue kit (Qiagen, Valencia, CA, USA) following manufacturer's instructions. DNA from IS was extracted according to the same process, after having re-suspended the parasite pellets in 200 µL PBS.

All DNA samples were genotyped using four microsatellite primers (MisatG4, MisatG9, MicBg5 and MicBg6) previously designed by screening for dinucleotide repeats

from a genomic bank of a *Tbb* strain (Koffi et al., 2007) and then labelled as *brucei* primers (BP). We also designed 14 new primers in the frame of this study. They were obtained from the genomic DNA of *Tbg* reference strain B4F303 (Koffi et al., 2007) and were labelled as *gambiense* primers (GP). These primers (name, repeated unit, sequence and annealing temperature) are described in the Supplementary file S1.

For BP, DNA samples were processed as previously described with a two-step amplification using a nested PCR (Kaboré et al., 2011). For GP, amplification conditions are described in the Supplementary file S1. After PCR amplification of microsatellite loci, alleles were routinely resolved on an 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. Amplification failures were noted in case of absence of amplification product or when the profile could not be interpreted.

Data analyses

All data sets were designed for Create (Coombs et al., 2008) and converted into the appropriate format for the different analyses.

Comparisons of amplification success between BS and IS or between BP and GP were undertaken with a Fisher exact test with R (R-Core-Team, 2018). Each test was performed within a homogeneous subsample: with BP only and with GP only for the BS/IS comparison; within BS or IS only for BP/GP comparisons. For these comparisons, we excluded uninterpretable profiles. This ended with four non-independent *p*-values. We thus corrected this four tests series with the Benjamini and Yekutieli procedure (BY) (Benjamini and Yekutieli, 2001) with R (command p.adjust). We also tested the proportions of loci that

were correctly amplified and those that presented an uninterpretable result with the same test and following the same steps as explained above.

A substantial proportion of homozygous profiles could be due to amplification failures of one allele in a heterozygous genotype (Kaboré et al., 2011). Given the amplification failures observed in our study, we used a one-sided Spearman's correlation test to test whether there is a positive relationship between the proportion of blank, of uninterpretable genotypes (among visible genotypes), of total missing profiles (blanks+uninterpretable) and the proportion of homozygous profiles (among loci with a clear genotype) found per stock.

We considered each cohort/focus or each cohort/village combinations as the units defining subsamples for analyzing local subpopulation structure.

To visualize the organization of stocks among each other's and in particular stocks from animals as compared to *Tbg*1 reference stocks, we built a Neighbor-Joining-Tree (NJTree) (Saitou and Nei, 1987) from a Cavalli-Sfoza and Edwards chord distance matrix (Cavalli-Sforza and Edwards, 1967), as recommended (Takezaki and Nei, 1996). We computed genetic distances with FreeNA (without INA correction) (Chapuis & Estoup, 2007), and constructed the tree with Mega 7 (Kumar et al., 2024). For this analysis, we kept only samples that had at least 70% of interpretable loci for *gambiense* and *brucei* primers to avoid empty lines in the matrix and to limit biases in genetic distance estimates.

Population genetic structure was evaluated through linkage disequilibrium (LD) tests and Wright's fixation indices (Wright, 1965). LD tests were performed by 10,000 randomizations of genotypes for each locus pair with the *G*-based test, which is the most powerful for combining tests across subsamples (De Meeûs et al., 2009). We undertook these tests with Fstat 2.9.4 (Goudet, 2003) an updated version of Fstat 1.2 (Goudet, 1995). There was as many non-independent tests as locus pairs (here 18×17/2=153 pairs). We thus had to correct this test series with the BY procedure.

For a hierarchy with three levels (individuals in subsample in total sample), three *F*-statistics can be defined: F_{IS}, which 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 (degree of 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 (one heterozygous class) to +1 (all individuals homozygous) and F_{ST} from 0 (all subsamples share the same allele frequencies) to +1 (all subsamples fixed for one or the other allele).

F-statistics were estimated with Weir and Cockerham's unbiased estimators (Weir and Cockerham, 1984) with Fstat 2.9.4: f for F_{IS}, θ for F_{ST}, and F for F_{IT}. These estimators vary similarly to their corresponding parameters except θ that can display negative values when subsample share allele frequencies that are more similar than expected under random sampling of the same subpopulation.

Significance of deviation from random mating or of subdivision was assessed with 10,000 permutations of alleles between individuals of the same subsample (for F_{IS}), of individuals between subsamples (for subdivision) and of alleles between individuals over all subsamples (for F_{IT}). The statistics used for these randomization tests were Weir and Cockerham's unbiased estimators for F_{IS} and F_{IT} and the *G* statistics for subdivision (Goudet et al., 1996) computed over loci, which is the most powerful combination procedure (De Meeûs et al., 2009). All these tests were undertaken with Fstat.

Confidence intervals at 95% (95%CI) of *F*-statistics were computed with the Jackknife standard error over subsamples or loci; or through 5000 bootstraps over loci. These were computed with Fstat. More details can be found elsewhere (De Meeûs et al., 2007).

According to Goudet et al.'s criterion (Goudet et al., 1994), the level at which hierarchically nested subsample are significantly subdivided can be assessed by comparing their F_{IS} (De Meeûs et al., 2006). Comparison between F_{IS} obtained within foci (ignoring villages) ($F_{IS\text{-focus}}$) and within villages ($F_{IS\text{-village}}$) was undertaken with a one-sided Wilcoxon signed rank test for paired data with Rcmdr, the pairing criterion being the locus. The alternative hypothesis was $F_{IS\text{-focus}} > F_{IS\text{-village}}$, meaning that village is a better suited subdivision, because ignoring those into foci produces a significant increase of F_{IS} (Wahlund effect).

According to theoretical results, F_{IS} is expected to display negative values in clones (Balloux et al., 2003), with a variation across loci explained by the level of polymorphism of these loci (e.g. through different mutation rates) (Séré et al., 2014). Indeed, in clonal populations with an infinite allele model of mutation (IAM), $F_{IS}=-(1-H_S)/H_S$ (Séré et al., 2014), When variation of F_{IS} across loci is strong and not explained by H_S, it was suggested that it may come from either very rare sex or presence of amplification problems (e.g. null alleles) (Balloux et al., 2003; De Meeûs et al., 2006; Séré et al., 2014). There is a way to discriminate between the two causes (rare sex versus amplification problems) (Séré et al., 2014). For this, we needed to compute the difference between $F_{ISexp}=-(1-H_S)/H_S$ and the observed one (F_{IS_obs}): $\Delta F_{IS}=F_{IS_exp}-F_{IS_obs}$. When $|\Delta FIS| \leq 0.05 \times |F_{IS_exp}|$ the two values are considered superimposed, i.e. total clonality without amplification problems. When superimposed points proportion across loci $p_{sup} \leq 30\%$, rare sex better explains the data, while $p_{sup} > 80\%$ reflects more amplification problems with total clonality or negligible amounts of sexual recombination (i.e. clonal rate $c \geq 0.9999$). This criterion, that we will call Séré's criterion, was

applied to each locus with H_S>0.5 as recommended (Séré *et al*., 2014). In pure clones, a positive correlation is expected between F_{IS} and H_S. This correlation was tested (unilateral test) with a Spearman's rank correlation test under Rcmdr.

We studied the correspondence between Cavalli-Sforza and Edwards chord distance matrices obtained with BP and GP with a Mantel test. Matrices were obtained as described above and Mantel test undertaken with Fstat. Indeed, in clonal populations, a strong correlation is expected between individual-based genetic distances of independent sets of markers (Tibayrenc et al., 1990; Telleria et al., 2013).

Isolation by distance test 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 $F_{ST_R}=a+b\times\mathrm{Ln}(D_G)$, and where a is a constant (intercept), $\mathrm{Ln}(D_G)$ is the natural logarithm of the geographic distance between two sites and b is the slope of the regression. Because clonal reproduction tends to reduce F_{ST} values (Balloux et a*l*., 2003; De Meeûs and Balloux, 2005; De Meeûs et al., 2006), Rousset's model cannot be used to estimate demographic parameters from the slope of isolation by distance regression. Since some subsamples are not contemporaneous, we undertook the regression with distances between contemporaneous subsamples only (i.e. same cohort). Weir and Cockerham estimator of F_{ST} and bootstrap (5000) confidence intervals were computed between each contemporaneous subsample pairs with FreeNA (Chapuis and Estoup, 2007) (without ENA correction) and then transformed into F_{ST_R} and regressed against $\mathrm{Ln}(D_G)$. Isolation by distance was considered significant if the 95%CI of the slope did not contain 0. In case of no significance, we also undertook a Mantel test with Fstat with the Cavalli-Sforza and Edward's chord distance matrix (Cavalli-Sforza and Edwards, 1967) because this test can be more powerful in some instances (Séré et al., 2017). This was done to avoid type II error, which is expected to be frequent for these tests (Séré et al 2014). We converted the bilateral

p-value of Fstat into a unilateral one (correlation was expected to be positive) as described in (Manangwa et al., 2019): in case of positive correlation, the *p*-value of Fstat was halved, otherwise the unilateral *p*-value was 1-half the bilateral one.

We finally used De Meeûs and Balloux's criteria (De Meeûs and Balloux, 2005) to detect strongly subdivided clones, where $F_{ST} \approx -F_{IS}/(1-F_{IS})$ and $F_{IT} \approx 0$. For these analyses, we only kept subpopulations with enough genotyped isolates from the same cohort. We thus only kept subsamples from January 2017 to March 2017 (Cohort C3).

Simulations

In order to check if some parameter estimates were valid, we undertook some simulations to explore the different parameters that may lead to the pattern observed in animal *Tb* samples from Côte d'Ivoire. We were particularly interested in the observed F_{IS}, the number of repeated genotypes, the shape of the NJTree, and the goodness of fit and/or direction of deviation from De Meeûs and Balloux's and Séré et al's criteria (see above). Simulations were undertaken with Easypop v. 2.0.1 (Balloux, 2001). Different parameter sets were explored until we obtained similar results to those observed for animal trypanosome data. All simulations started with maximum diversity and ran for 10,000 generations, with 18 loci as in the trypanosome data set. We used the population genetics analyses results to help limit the possible parameter sets (see below). At the end of each simulation, we sampled 24 subpopulations of 11 individuals (as the number of visible genotypes in Gozy 2017, in the real data set).

Results

The row data of microsatellite genotyping and TgsGP PCR results are given for each BS (n = 60), IS (n = 40) and reference stocks (n = 27) (127 DNA in total) in the

Supplementary file S1. Among the reference stocks, only the *Tbg*1 stocks were tested positive for the TgsGP PCR excepted P26F isolated from a pig in Cameroon.

Amplification failures

Comparisons of amplification success are presented in Figure 2 and 3. It can be seen that globally, BS provided significantly more amplification failures than IS and that BP offered better results than GP. The higher proportion of amplification successes were obtained with IS genotyped with BP, but still with important amplification failures (almost 30%).

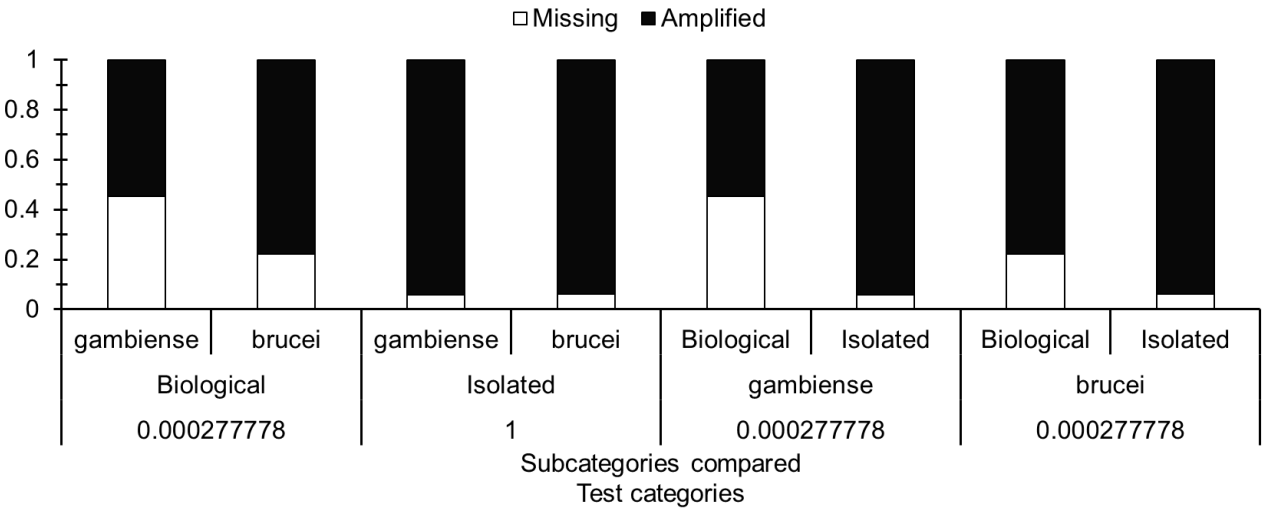

Figure 2: Proportions of correct amplifications (uninterpretable profiles ignored) between GP and BP within biological or isolated stock samples and between biological and isolated stock samples within GP or BP. The different *p-values* presented were corrected with the Benjamini and Yekutieli procedure for non-independent test series.

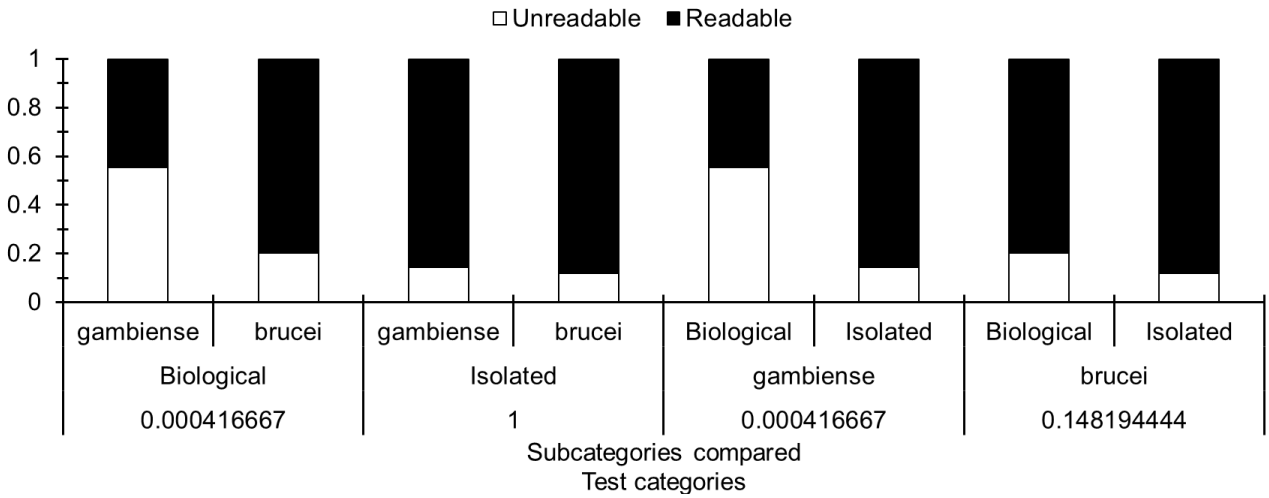

Figure 3: Proportions of readable amplifications (missing data ignored) between GP and BP within biological or isolated stock samples and between biological and isolated stock samples within GP and BP. The different *p-values* presented were corrected with the Benjamini and Yekutieli procedure for non-independent test series.

When focusing on reference strains only, we noticed that *Tbb* isolates were more easily amplified than other isolates, but the difference was weak (*p*-value=0.0333). To this respect, GP did not perform better than BP (*p*-value=0.1864).

The Spearman's correlation test showed that there was a strong and significant correlation between the proportion of blank genotypes found per locus and the proportion of uninterpretable profiles (ρ=0.6671, *p*-value=<0.0001). We also found a significant correlation between the proportion of missing data and the proportion of homozygous profiles (ρ=0.4314, *p*-value<0.0001), and between the proportion of uninterpretable profiles and homozygous ones (ρ=0.1762, *p*-value=0.0421). This means that the number of missing data is a fair measure of the quality of the DNA of the concerned individual: amplification failures increase the risk of uninterpretable results and the risk of dropouts of one of the two alleles in heterozygous profiles. This may also mean an underestimation of F_{IS}. Nevertheless, when keeping stocks with 70% of interpretable profiles, these correlations became non-significant (all *p*-values>0.23) and even negative between blank and homozygous profiles. Nevertheless, in this subset of data, the correlation between

uninterpretable and blank genotypes stayed strong and significant (ρ=0.4065, *p*-value=0.00085). Consequently, although the last correlation confirms the lowest quality of individuals with more blank and uninterpretable genotypes, this probably did not alter the estimate of heterozygote frequencies when keeping stocks with the highest amplification performances, as the one used in analyses described below.

NJTree analysis

Out of the 100 DNA samples from BS and IS, 34 displayed more than 70% amplification failures. Therefore, only 66 samples and the 27 reference stocks were considered for this analysis. The NJTree obtained with these samples is presented in Figure 4. All *Tbg*1 reference stocks took place in a distinct clade. All other samples including *Tbg*2 and *Tbb* reference stocks and both BS and IS, occurred in a scattered order in the tree. Surprisingly, the five TgsGP positive IS looked very distant from the *Tbg*1 clade. All came from the Vavoua focus, and two, from the same village, displayed the same multilocus genotype (same clone).

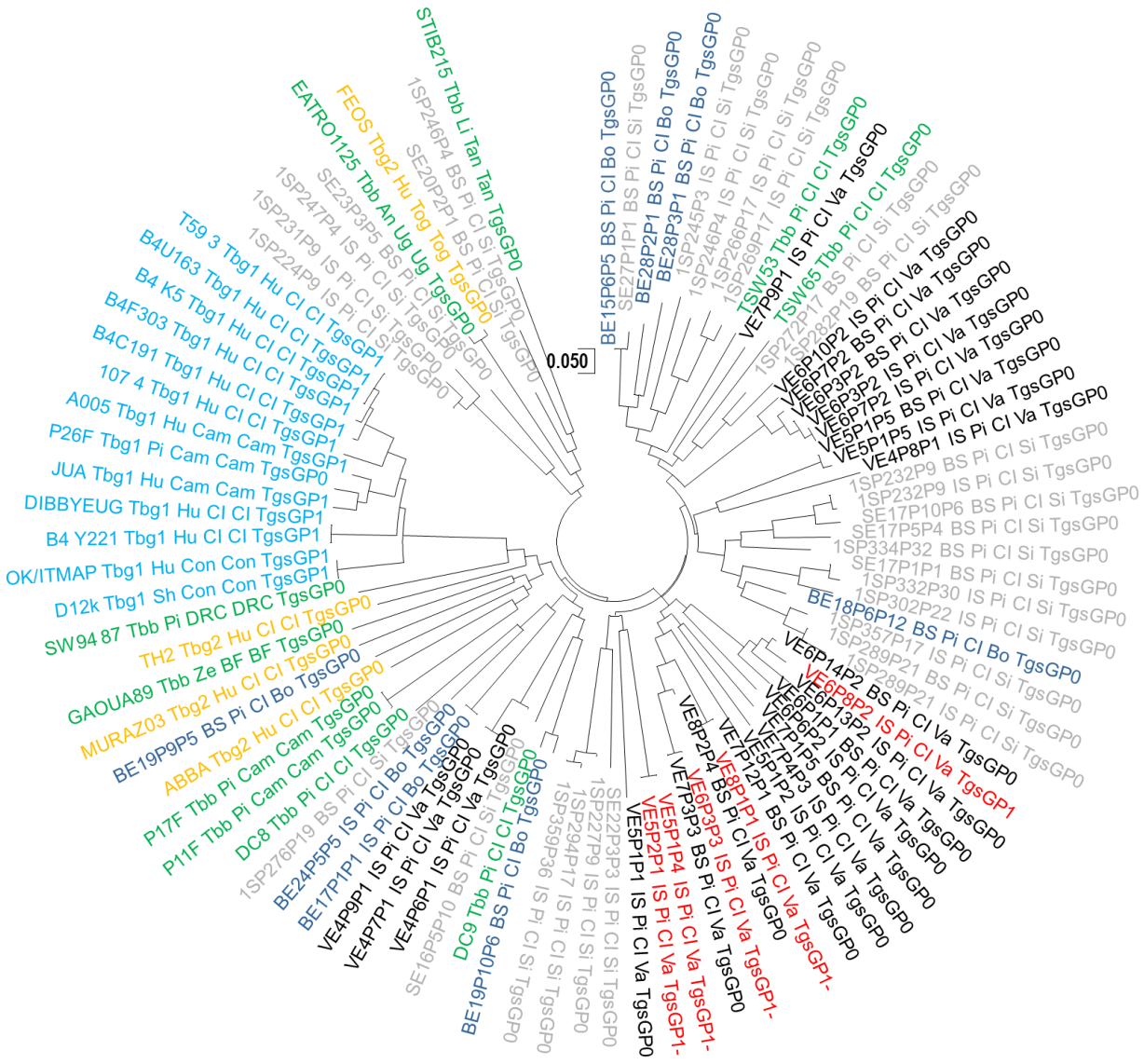

Figure 4: Neighbor-joining Tree (NJTree) of trypanosome stocks from a Cavalli-Sforza and Edwards chord distance matrix with the 66 biological and isolated stock samples that provided an interpretable profile for at least 70% of microsatellite primers among the 18 available, and 27 reference stocks. Reference strains of *Tbg*1 are in light blue, *Tbg*2 in gold, and *Tbb* in green. Stocks under study are in blue grey (Bonon = Bo), grey (Sinfra = Si) and black (Vavoua = Va) if TgsGP PCR negative and in red (only Vavoua) if TgsGP PCR positive. Hosts: Hu = Human, Ze = Zebu, An = Antelope, Pi = pig; BS = Biological Samples, IS = Isolated Samples; TgsGP1 = TgsGP PCR positive, TgsGP1- = TgsGP PCR weakly positive TgsGP0 = TgsGP PCR negative; Origin of stock: Cam = Cameroon, CI = Côte d'Ivoire, Con = Congo, DRC = Democratic Republic of Congo, Ug = Uganda, Tan =Tanzania, BF = Burkina-Faso, Tog = Togo.

Linkage disequilibrium

For Cohort/Focus combinations, there was a very high proportion of locus pairs in significant LD (93%) even after Benjamini and Yekutieli correction for which 80% of locus pairs stayed in significant LD. For Cohort/Village combinations, the proportion of locus pairs in significant LD was 85%, 46% of which stayed significant after Benjamini and Yekutieli correction. There is thus a global linkage disequilibrium across the whole genome of these *Tb* trypanosomes from pigs. Moreover, the proportions of significant LD tests in Cohort/Focus and in Cohort/Village combinations were significantly different (142/153 and 130/153 respectively, *p*-value=0.0441), which points out a possible Wahlund effect when pooling strains from different villages of the same focus.

Deviation from panmixia in foci or villages

There was a significant deviation from panmixia with heterozygote excesses in both foci and villages data sets. For foci, F_{IS}= -0.074 in bootstraps 95%CI=[-0.130, -0.022]. For villages F_{IS}= -0.16 in bootstraps 95%CI=[-0.221.. -0.102]. It thus appeared that F_{IS} was higher in foci than in villages and the difference was highly significant (*p*-value=0.0002). This translates into the fact that the real subpopulation unit corresponded more to villages than to foci. In line with the LD tests, in foci, an obvious Wahlund effect artificially increased the F_{IS} values at all loci. Finally, it also appeared that variation across loci was very weak (Figure 5), despite the small sample sizes and the substantial proportion of amplification failures. The correlation between F_{IS}-village and number of missing data (blanks genotypes) was low and not significant (ρ=0.103, *p*-value= 0.3415). It means that null alleles did not affect much F_{IS}. Within villages, Séré's criterion was never met since p_{sup}=0%. Nevertheless, the correlation between F_{IS} and H_S was positive and significant (ρ=0.620, *p*-value=0.0036). This

comes from the fact that genetic diversity, though quite high (H_S= 0.684), appeared not high enough for the level of F_{IS} observed to fit Séré et al's criterion.

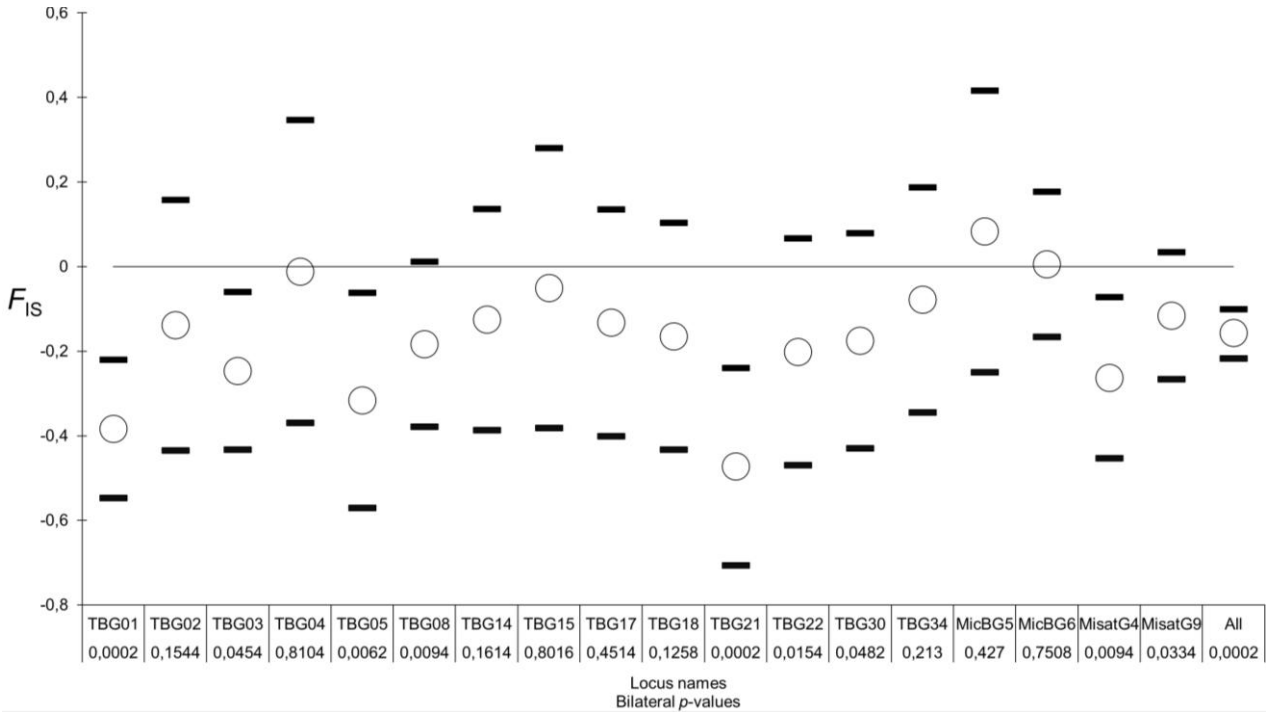

Figure 5: F_{IS} values observed in subsamples defined with cohort/village combinations (black dots) with their 95% confidence intervals (95%CI) (black dashes). The 95% CI of the average was obtained by bootstrap over loci and the 95% CI of each locus was obtained by Jackknife over populations. The two-sided *p*-values obtained after randomization of alleles between individuals in each subsample are indicated.

Mantel test

The correlation between BP and GP genetic distance matrices was positive and highly significant (ρ=0.23, *p*-value<0.0001), as expected in strongly clonal populations.

Isolation by distance and subdivision

Isolation by distance appeared not significant according to the slope (*b*=0.0058 in 95%CI=[-0.0103..0.0267]), even when considering subsamples sizes with at least 5 individuals or more. The Cavalli-Sforza and Edward's chord distance matrix-based Mantel test also provided a non-significant *p*-value=0.1865, with a weak correlation *r*=0.0834 and

determination coefficient R^2=0.007. Nevertheless, the average F_{ST} across contemporaneous subsample pairs, that could be computed for members of cohort 3, was F_{ST}≈0.108 with 95%CI=[0.086..0.130], and a highly significant *p*-value=0.0002. A substantial subdivision exists, but it appeared weakly connected with the spatial organization of sites. Indeed, trypanosome subpopulations from some remote sites were genetically closely related and some others from geographically nearby sites appeared genetically distant.

De Meeûs and Balloux's criteria

For members of Cohort 3, F_{IS}=-0.163 in bootstraps 95%CI=[-0.225..-0.105]. According to De Meeûs and Balloux's prediction, this would predict a F_{ST}=$-F_{IS}/(1-F_{IS})$=0.140 in bootstraps 95%CI=[0.085..0.174], which was obviously not significantly different from the F_{ST} computed above. Additionally, F_{IT}=-0.037 in bootstraps 95%CI=[-0.093..0.012] was not significantly different from 0 according to bootstraps and randomization tests (*p*-value=0.5072). These two observations are in line with a total clonality in a strongly subdivided population. If so, the number of immigrants entering a subpopulation per generation could be computed as (De Meeûs and Balloux, 2005) $Nm=-(1+F_{IS})/(4F_{IS})$=1.28 in 95%CI=[1..2]. The trypanosomes from pigs sampled in different regions of Côte d'Ivoire thus seemed to behave as a completely clonal and highly subdivided population. Given the F_{IS} range, quite far from -1, subpopulations probably displayed important sizes and then very weak immigration rates.

Simulations

During the exploratory phase, any simulation with a clonal rate c<1 failed to produce data sets in agreement with De Meeûs and Balloux's criterion. We then explored the case c=1 for various subpopulation numbers (n between 25 and 100), subpopulation sizes (N

between 1000 and 10000), immigration rates (*m* between 0.001 and 0.0001), mutation rates (*u* between 10^{-5} and 10^{-3}) and maximum number of alleles (*K* between 18 and 30). We then compared the shape of the NJTree obtained, the F_{IS}, F_{ST} and F_{IT} to those observed in the real animal trypanosome data set. The best fit was obtained for *n*=25, *N*=1000, *m*=0.001, *u*=0.001 and *K*=18. No significant difference was obtained between the values observed in the real dataset and those obtained in this simulation (Figure 6 and 7).

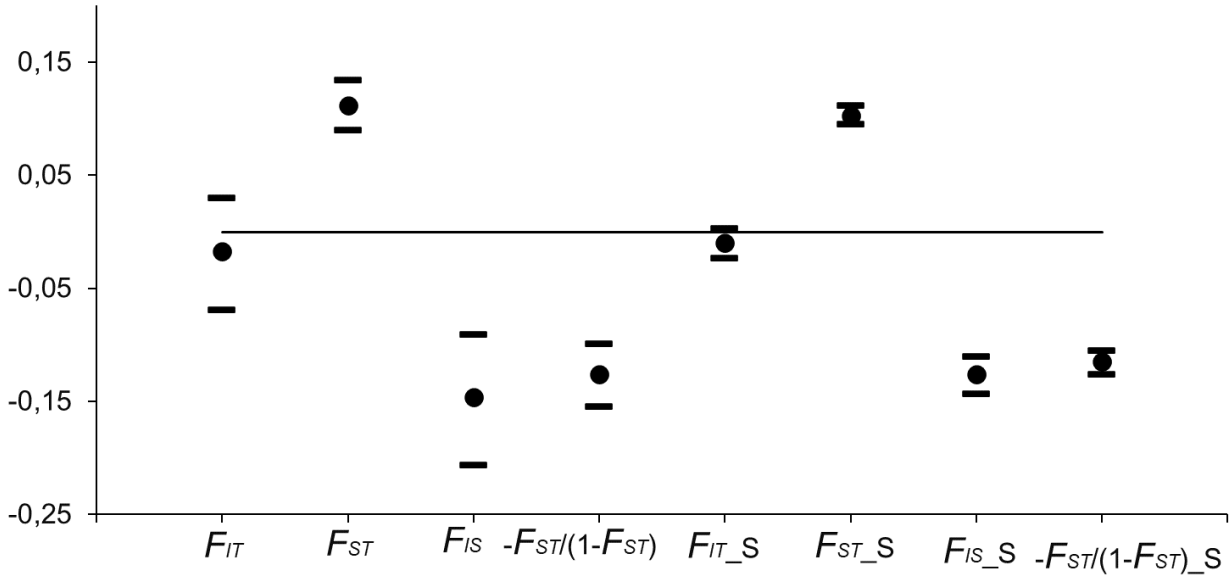

Figure 6: Comparison of genetic parameter values (F_{IS}, F_{ST} and F_{IT}) between the real values observed in trypanosomes from Côte d'Ivoire and the values obtained with our best simulation (_S). For each parameter, 95% confidence intervals (95%CI) obtained by bootstrap over loci are represented.

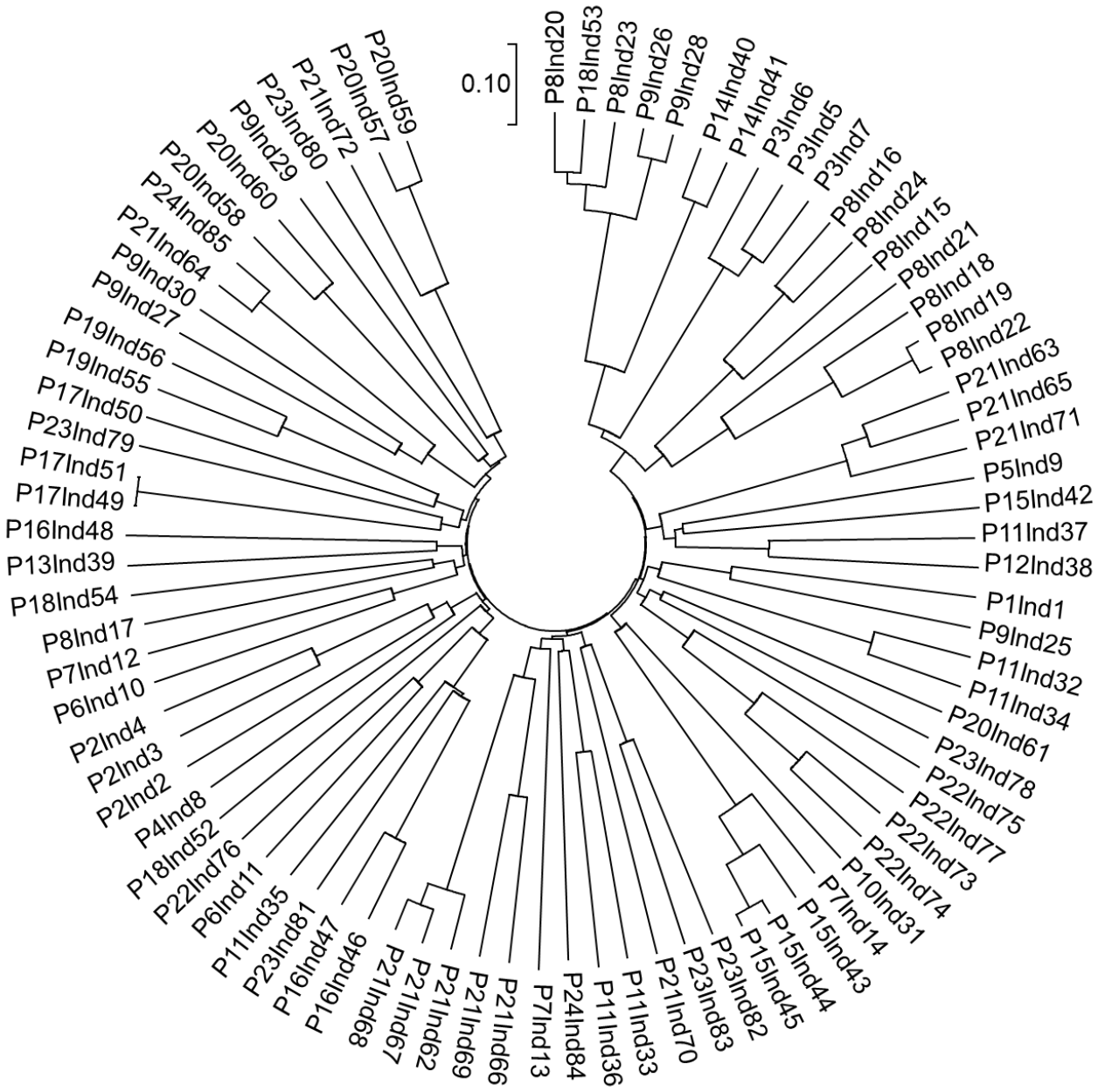

Figure 7: Cavalli-Sforza and Edward's chord distance based NJTree obtained with the simulation with the best fit to animal trypanosomes subsamples from Côte d'Ivoire.

Discussion

In Côte d'Ivoire, pigs appeared as able to play an important role in the transmission, the propagation and the spread of *Tb* s.l. parasites. In this context, the characterization of trypanosomes found in pigs could give a real indication of the genetic diversity of *Tb* s.l.

circulating in Ivoirian HAT foci, clarify the relationships between *Tbb* and *Tbg* and improve our comprehension of HAT epidemiology.

Before interpreting the population genetics approach results, we have to discuss the surprising results obtained with the TgsGP PCR. The five positive IS from Vavoua did not fit into the *Tbg*1 clade according to the NJTree (Figure 4). The specificity of TgsGP PCR was already questioned since no HAT case has been detected in this historical focus since 2011 (Traoré et al.,2021). In a recent study targeting the TgsGP gene to detect *Tbg* in tsetse flies in Uganda, the TgsGP primers were suspected to cross-react with DNA from an “unidentified source” after sequencing the DNA products (Cunningham et al., 2020). Although the TgsGP gene is an identification target for *Tbg*1 (Radwanska et al.,2002b; Gibson et al.,2010), similar genes have been identified in other *Tb* s.l. isolates (Felu et al., 2007). Indeed, according to Gibson et al. (2010), TgsGP-like genes can be found in some isolates of *Tbb*, *Tbr* and *Tbg*2. These genes, are closely similar to VSG genes. Together with the occurrence of a TgsGP PCR negative *Tbg*1 reference stock (P26F), this suggests that some VSG genes harbor sequences that are closely related to the sequences that are targeted by the TgsGP primers. Consequently, a few mutations may recurrently, though rarely, produce TgsGP positive results in *Tbb* lineages that may not be necessarily human serum resistant. Symmetrically, such mutations may reverse the process and produce apparent TgsGP negative profiles in *Tbg*1 lineages that may not necessarily become human serum sensitive. Sequencing such lineages at these genes would clarify such issues. Nevertheless, and given such observations, these primers are far from perfect and TgsGP PCR results should be interpreted with caution as they may lead to an over-estimation of the *Tbg* prevalence as recently observed (Ilboudo et al., 2025). The TgsGP PCR negative reference stock (P26F) previously identified as *Tbg*1 in several other studies and in the present one, was isolated from a pig in Cameroon and was used to argue for the existence of an animal reservoir of

Tbg (Nkinin et al., 2002; Simo et al., 2005; Simo et al., 2008). Further studies will be required to assess the human serum resistance of this particular strain and study the mechanisms involved.

TgsGP PCR and microsatellite genotyping results suggested that *Tbg*1 is not circulating at a detectable level in pigs in HAT foci in Côte d'Ivoire. This can be explained by the lack of sensitivity of single gene copy targets. Both TgsGP and microsatellite primers target a single DNA sequence copy and may not be sensitive enough. Given the low parasitaemia generally observed in *Tbg* infections in human patients (Deborggraeve and Büscher, 2012), it is thus possible that *Tbg* parasites could be even rarer when in an animal host and missed by these two methods.

Given the available data, it thus seemed that *Tb* s.l. is composed of several lineages that propagate clonally, with the rare occurrence of lineages that adapt to human hosts, as did *Tbg*1 10,000 years ago (Weir et al., 2016). We used this time as a standard, assuming a molecular clock for D_{CSE}, and using maximum distances found in each group of the NJTree in Figure 4. This allowed calculating that data were compatible with the existence of clonal lineages as old as 27,000 years over all genotyped strains (D_{CSE}=0.9), with slightly smaller divergences in Bonon and Sinfra (24,000 and 26,000 years respectively), and maximum divergences found in Vavoua. Average distances between *Tbg*2 strains and others appeared smaller (22,000 years), but still more ancient than between *Tbg*1 and other lineages. If true, this would mean that transfer from animals recurrently, though rarely, occurs, with a very successful one (Tbg1), which propagated across all Western and Central Africa.

There was convincing indications that the populations investigated here was highly or more likely absolutely clonal and strongly subdivided, according to De Meeûs and Balloux's criterion, but the data did not entirely fit with Séré et al's criterion (only the

correlation between H_S and F_{IS} matched). This apparent contradiction may come from a higher homoplasy and mutation rate as compared to what was explored in Séré et al. (Séré et al., 2014). Simulations suggested that the results observed in these *Tbb* samples was compatible with a big clonal population (N_{total}=25,000), composed of strongly isolated demes (immigration rate m=0.001), and that the markers used displayed very high mutation rates (u=0.001) with a substantial homoplasy (number of possible alleles K=18). A better fit could have been reached if we had explored subpopulations of variable sizes and variable values for u and K across the 18 loci to adjust to trypanosome data, but this would have been extremely time consuming and out of the scope of the present paper. Absolute clonality is in contradiction with the belief that *Tbb* extensively sexually recombine in the wild (Capewell et al., 2013; Echodu et al., 2015). Comparison with other data is difficult because we did not use the same genetic markers. In addition, we took care of working with contemporaneous subsamples, avoiding temporal Wahlund effects. Date of sampling was not controlled in the studies from the cited literature, with isolates that may extend over, at least, 10 generations (Capewell et al., 2013) or even 150 generations (Echodu et al., 2015). Moreover, the variation across loci was much smaller in our data than in other published datasets. Additionally, genetic diversity appeared much smaller (Capewell et al., 2013) or even very weak (Echodu et al., 2015) as compared to what we observed with our microsatellite loci. This issue will require further investigations to ascertain the exact reproductive strategy of *Tbb* in the different zones where it can be found, but our results suggest that *Tbb* may be much more clonal in the wild than previously speculated.

For the 18 microsatellite primers considered here, a high level of genetic diversity was found in *Tb* strains circulating in animals, as illustrated by the Figure 4 and the relatively high level of genetic diversity measured. However, the primers used for this study revealed a significant amount of amplification failures during genotyping. This is particularly critical

for BS, meaning that population genetics of trypanosomes appears much easier with IS (Figure 2 and 3). Applying primers designed from a particular species, subspecies or remote population to more or less genetically distant related taxa may generate amplification problems (De Meeûs et al., 2019). Stocks from pigs all appeared as *Tbb* and it is not surprising that BP amplified better these stocks than GP. Nevertheless, there was still important numbers of amplification problems, whatever the primer kind.

The BP where designed from the genomic sequences of clone TREU 927/4 (Koffi et al., 2007). Clone TREU 927/4 was derived from the isolate GPAL/KE/70/EATRO 1534 that originated from Kiboko, Kenya. It was isolated from a *Glossina pallidipes* fly in 1970. It has a degree of resistance to human serum, though it lacks the SRA gene that is characteristic of the human infective subspecies *Tbr* from East Africa (Gibson, 2012). It is thus possible that this clone is genetically very distant from the different stocks we used in the present study, hence explaining the important amplification failures met. The correlation between loci homozygosity and missing data (Kaboré et al., 2011), proves that the primers have difficulties to produce correct PCR products. This suggests that there is an actual need for the development of other markers, may be from cloned samples that will require being selected as most representative of the diversity met in all the *Tbb* stocks of the present study. The fact that some primers failed to amplify some samples confirms the results of other studies (Koffi et al., 2007; Morrison et al., 2009) , especially on field samples without isolation of parasites. The low quantity of DNA for some *Tb* genotypes can also favor amplification failures and may explain the lack of some genotypes due to the presence of unreadable alleles.

Isolation by distance seemed not to structure this trypanosome population. However, the weak connectivity between the foci, the clonal propagation that tend to bias F_{ST} towards lower values, together with small subsample sizes may explain why we did not find any

signature of isolation by distance. Wherever possible, further studies should include more stocks per village to gain in power for such kind of tests. Nevertheless, this may also be due to an almost complete isolation between villages, indicating a poor success of immigrant strains, as tsetse flies and humans are supposed to be able to cross such distances.

In the context of HAT elimination, it is crucial to develop more accurate tools to study the epidemiological role of potential animal reservoir for *Tbg*. Controlled experimental infections of domestic animals as recently performed with *Tbb* in pigs (Ilboudo et al., 2023) will be needed to evaluate the specificity and sensitivity of the currently available tools and their usefulness in assessing the role of animals in the maintenance and circulation of *Tbg*.

To conclude, we confirm the limited efficiency of the TgsGP marker in characterizing unambiguously *Tbg*1 strains in the wild. This is why it is currently still very difficult to assess the precise role of animal and/or human reservoirs in the epidemiology of sleeping sickness. Our results also suggest that the entire taxon *Tb* s.l. may be composed of different, sometimes highly divergent, clonal lineages, a few of which adapted to human hosts, among which *Tbg*1 appeared as the most successful one to date. Finally, in the foci studied, villages appeared to harbor very big and diverse populations of trypanosomes, which may not be a very good news for the sake of animal and maybe human health. Nevertheless, with almost no exchange between villages, propagation risks seem very limited.

Acknowledgements

The authors thank Solenn Stoeckel and Sophie Arnaud-Haond for their useful advice for the choice of parameters for the simulations. We acknowledge all the technicians from the HAT teams of Institut Pierre Richet of Bouaké, University Jean Lorougnon Guédé of Daloa and HAT National Elimination Program of Côte d'Ivoire.

Funding

The work presented here is part of a global project on the population genetics of clonal organisms financed by the ANR PRC Generic 2018 Clonix2D. We gratefully acknowledge the financial support of the Bill Gates Fondation and from the ARTS program (Research allowance for a thesis in the South) of the Institut de Recherche pour le Development (IRD).

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Author contributions

Martial Kassi N'Djetchi: genotyping, data collection, data analyses and manuscript redaction.

Thierry De Meeûs: supervision, data analyses, writing of the manuscript and design of figures.

Barkissa Mélika Traoré: data collection and manuscript correction.

Sophie Ravel: genotyping and manuscript correction.

Adeline Ségard: genotyping and manuscript correction.

Jacques Kaboré: supervision and manuscript correction.

Abe Allépo Innocent: data analyses and manuscript correction

Dramane Kaba: supervision and manuscript correction.

Djakaridja Berté: data analyses and manuscript correction.

Thomas Konan: data collection.

Bamoro Coulibaly: design of figures.

Pascal Grébaut: culture and strain isolation

Geraldine Bossard: culture and strain isolation.

Bruno Bucheton: supervision and manuscript correction.

Mathurin Koffi: supervision, data collection and manuscript correction.

Vincent Jamonneau: supervision, data collection and manuscript correction.

Data, scripts, code, and supplementary information availability

Raw data are available online: XXXXDOI of the webpage hosting the data. For data analyses, all R scripts used are inserted in the main text. For other analyses we used facilities without the need of scripts.

References

Balloux, F., 2001. EASYPOP (version 1.7): A computer program for population genetics simulations. J. Hered. 92, 301-302.

Balloux, F., Lehmann, L., De Meeûs, T., 2003. The population genetics of clonal and partially clonal diploids. Genetics 164, 1635-1644.

Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 29, 1165–1188.

Büscher, P., Cecchi, G., Jamonneau, V., Priotto, G., 2017. Human African trypanosomiasis. Lancet 390, 2397-2409.

Büscher, P., Bart, J.M., Boelaert, M., Bucheton, B., Cecchi, G., Chitnis, N., Courtin, D., Figueiredo, L.M., Franco, J.R., Grébaut, P., Hasker, E., Ilboudo, H., Jamonneau, V., Koffi, M., Lejon, V., MacLeod, A., Masumu, J., Matovu, E., Mattioli, R., Noyes, H., Picado, A., Rock, K.S., Rotureau, B., Simo, G., Thévenon, S., Trindade, S., Truc, P., Van Reet, N., 2018. Do cryptic

reservoirs threaten gambiense-sleeping sickness elimination? Trends Parasitol. 34, 197-207.

Capewell, P., Cooper, A., Duffy, C.W., Tait, A., Turner, C.M.R., Gibson, W., Mehlitz, D., MacLeod, A., 2013. Human and animal trypanosomes in Côte d'Ivoire form a single breeding population. Plos One 8, e67852.

Cavalli-Sforza, L.L., Edwards, A.W.F., 1967. Phylogenetic analysis: model and estimation procedures. Am. J. Hum. Genet. 19, 233-257.

Chapuis, M.P., Estoup, A., 2007. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24, 621-631.

Coombs, 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., Jamonneau, V., Kambire, R., Solano, P., 2010. Crise ivoirienne et rapatriés burkinabés: évaluation et vérification du risque de réémergence de la maladie du sommeil au Burkina Faso. Med Trop in press.

Cunningham, L.J., Lingley, J.K., Tirados, I., Esterhuizen, J., Opiyo, M., Mangwiro, C.T.N., Lehane, M.J., Torr, S.J., 2020. Evidence of the absence of human African trypanosomiasis in two northern districts of Uganda: Analyses of cattle, pigs and tsetse flies for the presence of *Trypanosoma brucei gambiense*. PLoS Negl. Trop. Dis. 14, e0007737.

De Meeûs, T., Balloux, F., 2005. F-statistics of clonal diploids structured in numerous demes. Mol. Ecol. 14, 2695-2702.

De Meeûs, T., Lehmann, L., Balloux, F., 2006. Molecular epidemiology of clonal diploids: A quick overview and a short DIY (do it yourself) notice. Infect. Genet. Evol. 6, 163-170.

De Meeûs, 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 Meeûs, 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 Bioinformatics 10, 443.

De Meeûs, T., Chan, C.T., Ludwig, J.M., Tsao, J.I., Patel, J., Bhagatwala, J., Beati, L., 2019. Deceptive combined effects of short allele dominance and stuttering: an example with *Ixodes scapularis*, the main vector of Lyme disease in the U.S.A. bioRxiv 622373, ver. 4 peer-reviewed and recommended by Peer Community In Evolutionary Biology, doi: https://doi.org/10.1101/622373.

Deborggraeve, S., Büscher, P., 2012. Recent progress in molecular diagnosis of sleeping sickness. Expert Rev Mol Diagn 12, 719-730.

Dje, N.N., Miezan, 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.

Duvallet, G., Stanghellini, A., Saccharin, C., Vivant J.. 1979. Human trypanosomiasis focus of Vavoua (Ivory Coast). A clinical, parasitological and sero-immunological survey. Med Trop. 39: 517–26. PMID: 231175.

Echodu, R., Sistrom, M., Bateta, R., Murilla, G., Okedi, L., Aksoy, S., Enyioha, C., Enyaru, J., Opiyo, E., Gibson, W., Caccone, A., 2015. Genetic diversity and population structure of *Trypanosoma brucei* in Uganda: implications for the

epidemiology of sleeping sickness and Nagana. PLoS Negl. Trop. Dis. 9, e000335

Felu, C., Pasture, J., Pays, E., Perez-Morga, D., 2007. Diagnostic potential of a conserved genomic rearrangement in the *Trypanosoma brucei gambiense*-specific TgsGP locus. American Journal of Tropical Medicine and Hygiene 76, 922–929.

Franco, J.R., Priotto, G., Paone, M., Cecchi, G., Ebeja, A.K., Simarro, P.P., Sankara, D., Metwally, S.B.A., Argaw, D.D., 2024. The elimination of human African trypanosomiasis: Monitoring progress towards the 2021–2030 WHO road map targets. PLoS Negl. Trop. Dis. 18, e0012111. https://doi.org/10.1371/journal.pntd.0012111.

Gibson, W. C, Marshall, T. F. De. C. & Godfrey, D. G,. 1980. Numerical analysis of enzyme polymorphism: a new approach to the epidemiology and taxonomy of trypanosomes of the subgenus Trypanozoon. *Advances in Parasitology* 18, 175-246.

Gibson, W.C., 1986. Will the real *Trypanosoma brucei gambiense* please stand up? Parasitology Today 2, 255–257.

Gibson, W., Nemetschke, L., Ndung'u, J., 2010. Conserved sequence of the TgsGP gene in Group 1 *Trypanosoma brucei gambiense*. Infect. Genet. Evol. 10, 453-458.

Gibson, W., 2012. The origins of the trypanosome genome strains *Trypanosoma brucei brucei* TREU 927, *T. b. gambiense* DAL 972, *T. vivax* Y486 and *T. congolense* IL3000. Parasit. Vect. 5.

Godfrey, D. G., Baker, R. D., RiCkman, L. R. & Mehlitz, D., 1990. Advances in Parasitology, 29, 1-39. JACQUARD, A. (1973). Cahiers d'Anthropologie et

d'Ecologie Humaine, 1, 11-124.

Godfrey, D.G., Catherine, M.S., Gibson, W.C, Mehlitz, D., Zillmann, U., 1987. Enzyme polymorphism and the identity of *Trypanosoma brucei gambiense*. *Parasitology*, 94, 337-34.

Goudet,J., De Meeûs, T., Day A.J., Gliddon C.J., 1994. The different levels of population structuring of dogwhelks, *Nucella lapillus*, along the south Devon coast. In: *Genetics and Evolution of Aquatic Organisms* (ed Beaumont A), pp. 81–95. Chapman and Hall, London.

Goudet, J., 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. J. Hered. 86, 485-486.

Goudet, J., Raymond, M., De Meeûs, T., Rousset, F., 1996. Testing differentiation in diploid populations. Genetics 144, 1933-1940.

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).

Hoare, C.B., 1972. The trypanosomes of mammals. A zoological monograph. Blackwell, Oxford.

Hide, G., Cattand, P., Le Ray, D.J., Barry, D., Tait, A., 1990. The identification of *Trypanosoma brucei* subspecies using repetitive DNA sequences. *Molecular and Biochemical Parasitology*, 39 (1990) 213-226.

Ilboudo, H., N'Djetchi, M.K., Kaboré, W.J., Kaboré, J., Traoré, B.M., Tahita, M.C., Ahouty, B.A., Deborggraeve, S., Eloiflin, R., Ségard, A., Bucheton, B., Koffi, M., Jamonneau, V., 2025. Evaluation of the AnTat A/B and LiTat A/B primers for the detection of Trypanosoma brucei gambiense. Exp. Parasitol. 271, 108929. https://doi.org/10.1016/j.exppara.

Ilboudo, K., Boulangé, A., Hounyèmè, R.E., Gimonneau, G., Kaboré, J., Belem, A.G.M., Desquesnes,M., Lejon, V., Koffi, M., Jamonneau, V., Thévenon, S., 2023. Performance of diagnostic tests for Trypanosoma brucei brucei in experimentally infected pigs. PLoS Negl. Trop. Dis. 17, e0011730. https://doi.org/10.1371/journal.pntd.0011730.

Jamonneau V., N'Guessan P., N'Dri L., Simarro P., Truc P., 2000. Exploration of the distribution of *Trypanosoma brucei* ssp. in West Africa, by multilocus enzyme electrophoresis. Ann Trop Med Parasitol 94:643–649.

Jamonneau, V., Ravel, S., Koffi, M., Kaba, D., Zeze, D.G., Ndri, L., Sane, B., Coulibaly, B., Cuny, G., Solano, P., 2004. Mixed infections of trypanosomes in tsetse and pigs and their epidemiological significance in a sleeping sickness focus of Cote d'Ivoire. Parasitology 129, 693-702.

Jamonneau, V., Truc, P., Grébaut, P., Herder, S., Ravel, S., Solano, P., De Meeus, T., 2019. *Trypanosoma brucei gambiense* Group 2: The Unusual Suspect. Trends Parasitol. 35, 983–995. https://doi.org/10.1016/j.pt.2019.09.002.

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 Côte d'Ivoire. Tropical Med. Int. Health 11, 136-143.

Kaba, D., N’Gouan, K.E., Djohan, V., Berté, D., Blé, S.L., Kouakou, L., Jamonneau, V., Rotureau, B., 2020. Diagnostic clinique et traitement de la trypanosomiase humaine africaine dans le contexte d’élimination. Asssociation des Anciens Élèves de l’Institut Pasteur 62, 49-56.

Kaba, D., Djohan, V., Berte, D., Bi Tra, D., Selby, R., Koffi Alain De Marie Kouadio, K.A.,

Coulibaly, B., Traore, G., Rayaisse, JB., Fauret, P., Jamonneau, V., Lingue, K., Solano, P., Torr, S.J., Courtin, F., 2021. Use of vector control to protect people from sleeping sickness in the focus of Bonon (Côte d'Ivoire).

Kaba, D., Koffi, M., Kouakou, L., N'Gouan, E.K., Djohan, V., Courtin, F., N'Djetchi, M.K., Coulibaly, B., Adingra, G.P., Berté, D., Ta, B.T.D., Koné, M., Traoré, B.M., Sutherland, S.A., Crump, R.E., Huang, C.-I., Madan, J., Bessell, P.R., Barreaux, A., Solano, P., Crowley, E.H., Rock, K.S., Jamonneau, V., 2023. Towards the sustainable elimination of gambiense human African trypanosomiasis in Côte d'Ivoire using an integrated approach. PLoS Negl. Trop. Dis. 17, e0011514. https://doi.org/10.1371/journal.pntd.0011514.

Kaboré, J., MacLeod, A., Jamonneau, V., Ilboudo, H., Duffy, C., Camara, M., Camara, O., Belem, A.M.G., Bucheton, B., De Meeûs, T., 2011. Population genetic structure of Guinea *Trypanosoma brucei gambiense* isolates according to host factors. Infect. Genet. Evol. 11, 1129-1135.

Kabore J, Camara O, Koffi M, Sanou D, Ilboudo H, Sakande H, Camara M, De Meeus T, Ravel S, Belem AMG, MacLeod A, Bucheton B, Jamonneau V, Thevenon S. 2018. Differences in pathogenicity and virulence of Trypanosoma brucei gambiense field isolates in experimentally infected Balb/C mice. Infect Genet Evol, 63, 269-276.

Kambire, R., Lingue, K., Courtin, F., Sidibe, I., Kiendrebeogo, D., N'Gouan, K.E., Ble, L., Kaba, D., Koffi, M., Solano, P., Bucheton, B., Jamonneau, V., 2012. Sleeping Sickness Surveillance in Cote D'ivoire and Burkina Faso. Parasite 19, 389-396.

Koffi, M., Solano, P., Barnabé, C., De Meeûs, T., Bucheton, B., N'Dri, L., Cuny, G., Jamonneau, V., 2007. Genetic characterisation of *Trypanosoma brucei ssp* by microsatellite typing: new perspectives for the molecular epidemiology of human African trypanosomosis. Infect. Genet. Evol. 7, 675-684.

Koffi, M., De Meeûs, T., Bucheton, B., Solano, P., Camara, M., Kaba, D., Cuny, G., Ayala, F.J., Jamonneau, V., 2000. Population genetics of *Trypanosoma brucei gambiense*, the agent of sleeping sickness in Western Africa. *Proc Natl Acad Sci* U S A, 106(1):209–214.

Koffi, M., De Meeûs, T., Séré, M., Bucheton, B., Simo, G., Njiokou, F., Salim, B., Kaboré, J., MacLeod, A., Camara, M., Solano, P., Belem, A.M.G., Jamonneau, V., 2015. Population Genetics and Reproductive Strategies of African Trypanosomes: Revisiting Available Published Data. PLoS Negl. Trop. Dis. 9, e0003985. https://doi.org/10.1371/journal.pntd.0003985.

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.

Koné, M., N’Gouan, E.K., Kaba, D., Koffi, M., Kouakou, L., N’Dri, L., Kouamé, C.M., Nanan, V.K., Tapé, G.A., Coulibaly, B., Courtin, F., Ahouty, B., Djohan, V., Bucheton, B., Solano, P., Büscher, P., Lejon, V., Jamonneau, V., 2020. The complex health seeking pathway of a human African trypanosomiasis patient in Côte d’Ivoire underlines the need of setting up passive surveillance systems. PLoS Negl. Trop. Dis. 14, e0008588. https://doi.org/10.1371/journal.pntd.0008588.

Kone, M., Kaba, D., Kabore, J., Thomas, L.F., Falzon, L.C., Koffi, M., Kouame, C.M., Ahouty, B., Compaore, C.F.A., N'Gouan, E.K., Solano, P., Fevre, E., Buscher, P., Lejon, V., Jamonneau, V., 2021. Passive surveillance of human African trypanosomiasis in Cote d'Ivoire: Understanding prevalence, clinical symptoms and signs, and diagnostic test characteristics. PLoS Negl Trop Dis 15, e0009656.

Kumar, S., Stecher, G., Suleski, M., Sanderford, M., Sharma, S, and Tamura, K., 2024.

Molecular Evolutionary Genetics Analysis version 12 for adaptive and green computing. *Molecular Biology and Evolution* 41: 1-9 https://doi.org/10.1093/molbev/msae263.

Laveissière, C., Couret, D., Kiénou, J.P., 1981. Lutte contre les glossines riveraines à l'aide de pièges biconiques imprégnés d'insecticide, en zone de savane humide : 4. Expérimentation à grande échelle.

Lejon, V., Lindner, A.K., Franco, J.R., 2025. Human African trypanosomiasis. The Lancet 405, 937–950. https://doi.org/10.1016/S0140-6736(25) 00107-2.

Manangwa, O., De Meeûs, 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.

Mehlitz, D.M., Zillman, U., Scott, C. and Godfrey, D.G. (1982) Epidemiological studies on the animal reservoir of Gambian sleeping sickness, III. Characterisation of Trypanozoon stocks by isoenzymes and sensitivity to human serum. Tropenmed. Parasitol. 33, 113-118.

Mehlitz, D., 1986. Le re´servoir animal de la maladie du sommeil à *Trypanosoma brucei gambiense*. Etudes et Synthèses de l'IEMVT. n.18, Maisons-Alfort: CIRAD-IEMVT.156 pp. ISBN 2-85985-127-5.

Mehlitz D, Molyneux DH., 2019. The elimination of *Trypanosoma brucei gambiense*? Challenges of reservoir hosts and transmission cycles: Expect the unexpected. Parasite Epidemiol Control. 6: e00113. https://doi.org/10.1016/j.parepi.2019.e00113 PMID: 31528738.

Morrison, L.J., Tait, A., McLellan, S., Sweeney, L., Turner, C.M.R., MacLeod, A., 2009. A

Major Genetic Locus in *Trypanosoma brucei* is a Determinant of Host Pathology. PLoS Negl. Trop. Dis. 3, e557. https://doi.org/10.1371/journal.pntd.0000557.

Moser, D.R., Cook, G.A., Ochs, D.E., Bailey, C.P., Mckane, M.R., Donelson, J.E., 1989. Detection of *Trypanosoma congolense* and *Trypanosoma brucei* subspecies by DNA amplification using the polymerase chain reaction. Parasitology 99, 57-66.

Murray, P.K., Murray, M., Wallace, M., Morrison Morrisson W.L., McIntyre, W.I.M., 1977. Trypanosomiasis in N'Dama and Zebu cattle. I. An experimental investigation of susceptibility to *Trypanosoma brucei, T congolense* and mixed infections. In : International Scientific Council for Trypanosomiasis Research and Control. 15th Meeting, Banjul, Gambia, 1977. Vol 110. Organization of African Unity/Scientific, Technical and Research Commission.

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 gambiens*. PLoS Negl. Trop. Dis. 11.

Nei, M., Chesser, R.K., 1983. Estimation of fixation indices and gene diversities. Ann. Hum. Genet. 47, 253-259.

Nkinin, S.W., Njiokou, F., Penchenier, L., Grébaut, P., Simo, G., Herder, S., 2002. Characterization of *Trypanosoma brucei* s.l. subspecies by isoenzymes in domestic pigs from the Fontem sleeping sickness focus of Cameroon. Acta Trop. 81, 225-232.

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.

Radwanska, M., Claes, F., Magez, S., Magnus, E., Perez-Morga, D., Pays, E., Buscher, P., 2002. Novel primer sequences for polymerase chain reaction-based detection of Trypanosoma brucei gambiense. American Journal of Tropical Medicine and Hygiene 67, 289–295.

Ravel, S., De Meeûs, T., Dujardin, J.P., Zeze, D.G., Gooding, R.H., Dusfour, 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.

Rousset, F., 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics 145, 1219-1228.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.

Séré, M., Kabore, J., Jamonneau, V., Belem, A.M.G., Ayala, F.J., De Meeûs, T., 2014. Null allele, allelic dropouts or rare sex detection in clonal organisms : simulations and application to real data sets of pathogenic microbes. Parasites and Vectors 7, art. 331.

Séré, M., Thévenon, S., Belem, A.M.G., De Meeûs, 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.e.a., 2010. The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. Int J Health Geogr; 1: 9:57. https://doi.org/10.1186/1476-072X-9-57 PMID: 21040555.

Simo, G., Herder, S., Njiokou, F., Asonganyi, T., Tilley, A., Cuny, G., 2005. *Trypanosoma*

brucei s.l.: Characterisation of stocks from Central Africa by PCR analysis of mobile genetic elements. Exp. Parasitol. 110, 353-362.

Simo, G., Cuny, G., Demonchy, R., Herder, S., 2008. *Trypanosoma brucei gambiense*: study of population genetic structure of Central African stocks using amplified fragment length polymorphism (AFLP). Exp. Parasitol. 118, 172-180.

Stevens,J. R., Laniiam, S. M., Allingiiam, R. & Gashumba,J. K., 1992. *Annals of Tropical Medicine and Parasitology*, 86, 9-28.

Tibayrenc, M., Kjellberg, F., Ayala, F.J., 1990. A clonal theory of parasitic protozoa: the population structure of Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas and Trypanosoma, and its medical and taxonomical consequences. Proc. Natl. Acad. Sci. USA 87, 2414–2418.

Tait, A., Eldirdiri, A., Le Ray. D.,1984. Enzyme variation in *Trypanosoma brucei* spp. Parasitology 89. 311-326.

Takezaki, N., Nei, M., 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144, 389-399.

Telleria, J., Barnabé, C., F.J. Ayala, F.J., Tibayrenc,M., 2013. Phylogenetic character mapping of RADES Probing, a new marker for exploring the clonal evolution of expressed coding sequences in Trypanosoma cruzi, the agent of Chagas disease. *Infection, Genetics and Evolution*, 19, 287–291.

Traoré, BM, Koffi, M., N'Djetchi, KM., Kaba, D., Kaboré, J., Ilboudo, H., Ahouty, A.B, Koné, M., Coulibaly, B., Konan, T., Segard, A., Kouakou, L., De Meeûs, T., Ravel, S., Solano, P., Bart, JM., Jamonneau, V., 2021. Free-ranging pigs identified as a multi reservoir of *Trypanosoma brucei* and *Trypanosoma congolense* in the Vavoua area, a historical sleeping sickness focus of Côte d'Ivoire. PLoS Negl Trop Dis 15(12): e0010036. https://doi.org/10.1371/journal.pntd.0010036.

Truc, P., Mathieu Daudé, F., Tibayrenc, M., 1991. Multilocus isozyme identification of Trypanosoma brucei stocks isolated in Central Africa: evidence for an animal reservoir of sleeping sickness in Congo. *Acta Tropica*, 49, 127-135.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358-1370.

Weir, W., Capewell, P., Foth, B., Clucas, C., Pountain, A., Steketee, P., Veitch, N., Koffi, M., De Meeûs, T., Kaboré, J., Camara, M., Cooper, A., Tait, A., Jamonneau, V., Bucheton, B. Berriman, M., MacLeod, A., 2016. Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *Genomics and evolutionary biology Microbiology and infectious disease.* eLife 2015;5:e11473. DOI: 10.7554/eLife.11473.

WHO, 2013. Control and surveillance of human African trypanosomiasis: report of a WHO expert committee, World Health Organization & WHO Expert Committee on the Control and Surveillance of Human African Trypanosomiasis. World Health Organization. http://www.who.int/iris/handle/10665/95732 Geneva, Switzerland. Wright, S., 1965. The interpretation of population structure by F-statistics with special regard to system of mating. Evolution 19, 395-420.