Published in final edited form as:
Insect Mol Biol. 2001 August ; 10(4): 387–395.

Genetic differentiation of *Glossina morsitans centralis* populations

E. S. Krafsur, **M. A. Endsley**, **D. L. Wohlford**, **N. T. Griffiths**[1], and **R. Allsopp**[2]
Department of Entomology, Iowa State University, Ames, IA, USA

Abstract

Variation at mitochondrial and microsatellite loci was used to study the breeding and dispersal structure of *Glossina morsitans centralis*, in six natural populations from Botswana, the Caprivi Strip (Namibia), Zambia, and in a laboratory culture derived from Singida, Tanzania. Only seven mitochondrial haplotypes were found. Mean diversity averaged over the six natural populations was 0.216 ± 0.085. The fixation index F_{ST} = 0.866 indicated a high degree of genetic differentiation among populations. Fifty-three alleles were detected among six microsatellite loci and six natural populations. Mean microsatellite diversity was 0.702 ± 0.091. Depending on the estimating model used, fixation indices varied from 0.15 to 0.225 confirming that *G. m. centralis* populations are strongly subdivided. For all F_{ST} estimates, positive correlations were detected between pair-wise genetic distance measures and geographical distances. The difference in fixation indices estimated from mitochondrial or nuclear loci was explained by the greater sensitivity of mitochondrial genomes to genetic drift. Population differentiation can be explained by genetic drift and the subsequent recovery of extant populations from small, discontinuous populations. These data confirm genetically the collapse and retreat of *G. m. centralis* populations caused by the rinderpest epizootic of the late 19th and early 20th centuries.

Keywords

tsetse flies; *Glossina morsitans centralis*; gene flow; mitochondrial variation; microsatellite variation

Introduction

Tsetse flies (Diptera: Glossinidae) are distributed discontinuously, have low reproductive rates, and low population densities when compared with mosquitoes and synanthropic dipteran species (Ford, 1971; Rogers & Randolph, 1985). Therefore, it may be expected that tsetse populations would show pronounced degrees of genetic differentiation as a consequence of drift at selectively neutral loci unless dispersal rates were substantial. Methods of population genetics can provide estimates of spatial and temporal dispersal, but reports on tsetse population genetics are few. Most earlier work on tsetse population genetics, reviewed by Gooding (1992), was based on comparing independently derived laboratory cultures and few natural populations were studied directly. Eleven natural *G. pallidipes* Austen populations from Kenya, Zambia, Zimbabwe and Mozambique, however, showed remarkably high levels of genetic differentiation at allozyme loci (Krafsur *et al.*, 1997) and mitochondrial loci (Krafsur & Wohlford, 1999). Patterns of mitochondrial variation were examined also in *G. m.*

Correspondence: E. S. Krafsur, Department of Entomology, 403 Science 2, Iowa State University, Ames, IA 50011-3222, USA. Tel.: 515 294 8127; fax: 515 294 5957; e-mail:ekrafsur@iastate.edu.
[1]Formerly: Department of Biology, Imperial College at Silwood Park, Ascot, Berks SL5 0BH, UK.
[2]Formerly: Tsetse Control Division, Ministry of Agriculture, Botswana. Present address: Natural Resources Institute, University of Greenwich, Chatham Maritime, Chatham, Kent ME4 4TB, UK.

morsitans (Wohlford *et al.*, 1999) and high levels of population differentiation were observed. Solano *et al*. (2000) recorded in 201 *G. palpalis gambiensis* a high level of genetic differentiation at a microgeographic scale by scoring genotypic frequencies at three microsatellite loci. Two years of dry season samples suggested that flies accumulated in refugia from at least two gene pools.

Glossina morsitans s.l. is the most widespread tsetse. Three allopatric subspecies are known. We are concerned here only with *G. morsitans centralis* (Machado). A survey of allozyme variation in *G. m. centralis* showed 32% of the loci were polymorphic, and the mean heterozygosity was only 6.6%, a rather low value for insects (Krafsur & Griffiths, 1997). A rinderpest pandemic that began in 1887 all but eliminated *G. morsitans s.l.* mammalian hosts in much of Southern Africa. Resident entomologists mapped refugia and subsequent spread of *G. morsitans* in a portion of its range (reviewed in Ford, 1971). The present range of *G. m. centralis* encompasses foci in Angola, Botswana, Burundi, Namibia, Rwanda, Tanzania, Uganda, Zaire and Zambia (Jordan, 1993). We have samples from Botswana, Namibia, Tanzania and Zambia. We ask, are the populations represented by these samples freely interbreeding or are they derived from small, isolated populations? When populations become small and discontinuous, drift will cause gene frequencies to diverge in proportion to numbers and time since divergence. Natural selection also may cause genetic differentiation. The forces of drift and selection, however, are compensated by migration and so it is appropriate to ask, what levels of exchange obtain among the sampled populations? In other words, how much dispersal has there been among populations?

The methods chosen to evaluate gene flow were to measure and partition mitochondrial DNA (mtDNA) and microsatellite variation within and among populations. mtDNA is useful in population analysis because it is single copy, non-recombining, maternally inherited, and is subject to a high mutation rate. Typically, among-population variation is greater than within-population variation (Avise, 1994). Microsatellites are tandem nucleotide repeats that are both abundant and well distributed in the nuclear genome of numerous mammals (Jarne & Lagoda, 1996) and in *Drosophila* species (Noor *et al*., 2000). Some microsatellites are composed of di- and trinucleotide repeat arrays. Alleles vary in repeat number and hence in molecular weight and are codominant, allowing for the unambiguous detection of heterozygotes on gels or by automated methods.

Results

Mitochondrial loci

Seven populations of fourteen to 100 flies each were examined (Table 1). Six alleles were found at *12S* and five alleles at *16S2*. Only seven allelic combinations, or haplotypes, were detected.

Of 100 samples from the four Okavango delta traps, ninety-eight showed the same haplotype and were therefore considered to form a single population and were pooled. Chi-squared tests of homogeneity indicated highly significant differences in haplotype frequencies among the seven populations ($\chi_{(36)}^2 = 795.3$, $P << 0.001$). Unbiased estimates of diversity in each population, h_e, ranged from 0 to 0.434 with a mean over samples of 0.185 ± 0.078 (0.216 ± 0.085 for the six natural populations). These estimate the probability that any two randomly chosen flies in a population will have different haplotypes. Diversity among the pooled haplotypes was 0.543 ± 0.156.

F_{ST} measures the correlation of alleles within populations relative to the total. The estimate over the seven populations was $F_{ST} = 0.75$. By using Weir & Cockerham's (1984) method, $F_{ST} = 0.806 \pm 0.067$. Both measures suggest near fixation of haplotypes. F_{ST} estimates for

each pair-wise sample (Table 2) show that Okavango, Kwondo and Lupala populations were homogeneous, as were the Chunga and Katima populations. The pair-wise F_{ST}s, when converted to an equivalent number of female migrants, suggest very low rates (less than one fly per generation) of exchange between seventeen of twenty-one population pairs (Table 2). Dispersal among the remaining four populations was virtually unrestricted.

The foregoing pattern suggests the following grouping of populations for analysis of variance: Singida; samples from Zambia (Kasanka, Chunga, Katima); and samples from Namibia and Botswana (Lupala, Kwondo, Okavango). Analysis of variance (Table 3) on the haplotype frequencies indicated that 74.5% of the total variance lay between the foregoing three groups. Twelve per cent of the total variance lay among populations within groups and only 13.4% lay within populations. The associated fixation indices are F_{SG}, which estimates the probability that two randomly chosen flies within a population in a group have the same haplotypes relative to the probability of any two flies from the group as a whole. F_{GT} estimates the chances that two randomly chosen flies in a group have the same haplotypes relative to the total. F_{ST} took the value 0.866, a very high among-population value. It predicts an equivalent dispersal rate of one female tsetse every thirteen generations.

Microsatellite loci

Sample sizes were twenty-four flies. Allelic frequencies are presented in the annex. Fifty-three alleles were recognized over the six loci for a mean of 8.83 ± 3.71 per locus. Five 'private' alleles were detected (alleles confined to a single sample, 9.4 ± 4.0% of the total), these being found in the Chunga and Kwondo populations. Gene diversities averaged over the six loci varied from 49.4% at Kasanka to 69.1% at Lupala (Table 4). Each sample showed more heterozygotes than expected by Hardy–Weinberg criteria. Allele frequencies differed greatly ($P < 0.001$ at each locus) among the seven samples when tested by contingency ($\sum\chi_{[282]}^2 = 1.804$, $P << 0.001$). Averaged over populations, single locus heterozygosities (diversities) varied from 35% to 90% with a mean 70.2 ± 9.1% (Table 5).

Tests for mutation-drift equilibrium

Equilibrium populations show a relationship between allele diversity and heterozygosity that depends on the mutation rate and effective population size. Populations that have undergone recent, large changes in population size may not be in equilibrium between mutation and drift because mutation rates tend to lag changes in population magnitudes. Populations at equilibrium will have equal probabilities of showing a gene diversity excess or deficit and these can be tested by using the two-tailed Wilcoxon signed rank test (Cornuet & Luikart, 1996; Piry *et al.*, 1999). The two-phased mutation model incorporating 70% one-step mutation model and 30% infinite allele model was used to compute the distribution of the gene diversity expected from the observed number of alleles, given the sample size under the assumption of mutation-drift equilibrium. Results of the computations indicated that no population departed significantly from expectation ($P > 0.16$ at the most deviant locus).

F statistics

F_{IS} measures departures from random mating within populations and the estimates indicated surpluses of heterozygotes at five of six loci (Table 5). The mean estimate over loci $F_{IS} = -0.119 \pm 0.038$ was greater than the expected *c.* –0.027.

F_{ST} measures departures from random mating among populations; its mean over loci was 0.186 ± 0.035 (Table 5). This statistic estimates *c*. 1.1 migrants (both sexes) per generation. Pair-wise estimates of F_{ST} and *Nm* are presented in Table 6 and show dispersal rates that vary from *c*. 1–2 reproducing flies per generation. Only Singida, a closed laboratory strain, showed values less than unity, an unsurprising result. The mean frequency of private alleles estimated F_{ST} =

0.183 by Slatkin's method, which suggests 1.1 migrants per generation, half of which would be female assuming the sexes disperse at the same average rate.

Analysis of variance assigned only 7% of the variance among groups and 10.6% among populations within groups. Most variance, 82.3%, lay within populations (Table 7). The associated *F*-statistics suggest average levels of gene flow greater than unity, sufficient to prevent further genetic differentiation by pure drift.

Isolation by distance

Pair-wise F_{ST} were regressed on the corresponding pair-wise $\log_e$ distances to determine if there was a relationship between genetic distances and geographical distances. The Mantel test (Sokal & Rohlf, 1995) was used to test the statistical significance of the association. The procedure was carried out by using the FORTRAN program MANTEL written by W. C. Black IV (personal communication). Mitochondrial F_{ST} as a genetic distance measure was correlated with $\log_e$ geographical distance: $r = 0.72$, Mantel $P = \approx 0.001$. Pair-wise F_{ST} for microsatellite loci was correlated with geographical distance: $r = 0.65$, Mantel $P \approx 0.014$. Pair-wise estimates of F_{ST} for both mitochondrial haplotypes and microsatellite loci were compared by using the Mantel test and found to be correlated, $r = 0.58$, $P \approx 0.001$. Arithmetic plots of the pair-wise data (Fig. 1) show weak relationships between F_{ST} and distances greater than 400 km.

Discussion

Gene diversities

Comparatively little haplotype diversity was detected in *G. m. centralis*. Only seven haplotypes were detected and the mean diversity measured for the pooled data was 0.543. Twenty-six haplotypes were recorded in seven *G. m. submorsitans* populations for which pooled $h_e = 0.544$ (Krafsur *et al.*, 2000), twenty-three haplotypes in six *G. m. morsitans* populations for which pooled $h_e = 0.895$ (Wohlford *et al.*, 1999), and twenty-three haplotypes in twenty-one *G. pallidipes* populations for which $h_e = 0.630$ (Krafsur & Wohlford, 1999). These data suggest that *G. m. centralis* was reduced by the bottleneck to a lower density than the other taxa.

Microsatellite diversities in *G. m. centralis*, on the other hand, were much more substantial than mitochondrial diversities. Fifty-three alleles were recorded and the mean diversity (heterozygosity) was 0.702 ± 0.091. A comparable measure in *G. m. morsitans* was 0.558 ± 0.077 averaged over twelve loci (Baker & Krafsur, 2001). We found no evidence for null alleles at the six loci. Five private alleles were detected, testifying to restricted gene flow.

Among mammals, mutation rates at microsatellite loci tend to be high, of the order of 10^{-3}–10^{-5}. Although no single model accurately describes mutation dynamics, the step-wise model of mutation (SMM) probably is most appropriate for microsatellites, and leads to a limited number of allelic states and saturation (homoplasy), because length polymorphism is limited and each new mutation is dependent on the previous allelic state (Jarne & Lagoda, 1996; Gaggiotti *et al.*, 1999). In *Drosophila*, however, much lesser mutation rates obtain (Schug *et al.*, 1997). *Drosophila* and *Glossina* are phylogenetically related and may show similar mutation dynamics. The issue is important because high mutation rates and homoplasy tend to diminish estimates of F_{ST}, thereby overestimating rates of gene flow.

Gene flow

Only diploid loci can estimate gene flow within populations and the relevant statistic is F_{IS}, which indicated an excess of heterozygotes. It is more usual to observe a deficiency of heterozygotes, because of the Wahlund effect or the occurrence of null alleles. After correcting for type I errors in multiple comparisons by using sequential Bonferroni tests, only two of

forty-two comparisons (six loci × seven populations) showed a significant deviation ($P < 0.05$) from the expected number of heterozygotes. Moreover, tests for mutation-drift disequilibrium were negative, leading to the conclusion that populations are now probably close to mutation-drift equilibrium after having experienced a severe bottleneck and subsequent expansion in density and geographical range.

The mitochondrial data indicate substantial differentiation among the sampled populations, a consequence of geographical isolation and genetic drift. This is not altogether surprising because drift is a potent force in the small and geographically fragmented populations characteristic of *G. m. centralis*. Thus dispersal of reproducing females per generation, estimated from a mean $F_{ST} = 0.866$, was 0.08, or about one female every thirteen generations. Pair-wise population estimates of F_{ST} from mitochondrial and nuclear genes were congruent when plotted against the geographical distances between them.

Microsatellite allele frequencies provided independent estimates of F_{ST} that closely agreed for a mean $F_{ST} = 0.186$, equivalent to about 1.1 reproducing migrants per generation. Half are female, or 0.55 migrants, *c*. sevenfold the rate estimated by mitochondrial genes. Estimates of genetic differentiation based on mitochondrial and nuclear loci were also independent in *Drosophila* (Hale & Singh, 1991). How may we reconcile the differences between them? A migrant, gravid female can introduce four copies of a nuclear gene but only one copy of her mitochondrial genome. A migrant male carries two copies of nuclear genes but does not pass on his mitochondrial genome at all. Thus, the effective population size of mitochondrial genes is about one-quarter that of nuclear genes and the effects of genetic drift thereby become more pronounced (Avise, 1994; Hale & Singh, 1991). In principle, female tsetse flies may be less likely to disperse and reproduce than the males, although experiments do not support this idea (e.g. Rogers, 1977; Vale *et al.*, 1984). A further cause of differences in F_{ST} estimates between mitochondrial and microsatellite genes is the mutation dynamics of microsatellites. A high mutation rate and step-wise mutation acts to homogenize allele frequencies among populations thereby reducing F_{ST} estimates.

The foregoing estimates of gene flow need to be examined in the light of ecological research that indicates that *morsitans* group tsetse are highly vagile; *G. m. morsitans* may disperse over two dimensions at 230 m daily (Rogers, 1977) and we have no reason to think *G. m. centralis* would differ very much. Great vagility can rapidly minimize genetic differences between populations; in principle, numerically little exchange between populations prevents differentiation at selectively neutral loci – about one reproducing migrant per generation is the 'critical' level (Wright, 1978). What then of the substantial degree of genetic differentiation in *G. m. centralis*?

History provides the answer. The rinderpest panzootic of the last century is said to have reduced tsetse vertebrate host populations by *c*. 90% (Ford, 1971). Relict tsetse populations were confined to small foci from which they have subsequently spread. Those foci were recorded by resident entomologists. Tsetse control measures applied since then may have caused further loss of diversity. The history of insecticidal and other control measures applied to Botswanan and Namibian populations could explain the paucity of mitochondrial variation in the currently expanding populations there. Only two haplotypes were found among 174 tsetse examined, and the predominant one was 98.3% of the total. It seems likely that *G. morsitans* in Botswana and Namibia consist of fragmented populations descended from a very small ancestral population. It is interesting, however, that nuclear gene diversity was no less in the Botswana samples than elsewhere.

Estimates of gene flow are useful in the context of area-wide control measures. Use of the sterile insect technique (SIT) and its modifications is predicated upon maximizing matings

between sterile and wild flies. Estimates of gene flow provide an index of prevailing dispersal rates of wild flies. High rates of dispersal among wild flies could compromise the effectiveness of SIT, but the genetic data offered here suggests that the vagility of reproductive tsetse flies is not particularly great.

Experimental procedures

Sampling

Sampling locations are indicated in Fig. 2. The Zambia samples were taken in Kasanka, in the Kasanka National Park (*c.* 30°08′ E, 12°20′ S), and at Chunga, in the Kafue National Park (*c*. 25°55′ E, 15°05′ S). Tsetse were taken 67 km south of Katima Mulilo, in the Caprivi strip in Namibia (*c*. 24°00′ E, 18°00′ S). Lupala is an island in the River Linyanti in Mamili National Park of Namibia (*c*. 25°00′ E, 17°80′ S). The Botswana samples were from Kwando (*c*. 23°45′ E, 18°20′ S) and an arc of four traps in the Okavango delta (*c*. 23°31′ E, 19°15′ S) that varied from 15 to 54 km apart. The Kwondo site was 112 km north-east of the Okavango. In addition, we examined a laboratory culture maintained at the International Laboratory for Research on Infectious Diseases, Nairobi, Kenya. It is derived from flies collected at Singida, Tanzania (*c*. 35° E, 5° S) in 1969 (Moloo *et al.*, 1998). Loss of diversity by drift is not severe because successful tsetse husbandry demands high reproductive success from each female if a cultured population is to continue. Katima and Kwondo tsetse were treated to insecticidal ground sprays in the 1970s and 80s. Aerial sprays were applied to the Okavango for 20 years, beginning in 1972. The treated area varied from 2000 to 8000 km^2 annually (77 920 km^2 in total). Targets (Vale *et al.*, 1988) have been deployed since 1992.

Tsetse flies were collected in Zambia and Namibia by using Epsilon F3 cloth traps. The traps had pthalogen blue exteriors and black interiors and were baited with acetone, phenol and octenol (Torr *et al.*, 1989). Botswana flies were sampled by using Epsilon traps baited with octenol and methyl ethyl ketone in separate low-density polythene bottles. The flies were killed, frozen in liquid nitrogen or preserved in 80–90% ethanol, and shipped to Ames.

DNA extraction, primers and PCR amplification

A CTAB (hexadecyltrimethylammonium bromide) DNA extraction method (Shahjahan *et al.*, 1995) was used as described (Krafsur & Wohlford, 1999; Krafsur *et al.*, 2000).

To examine mitochondrial variation, primers for the 12S ribosomal RNA locus were SR-J-14233 and SR-N-14588. For the 16S2 ribosomal RNA locus, primers were N1-J-12585 and LR-N-12866 (Simon *et al.*, 1994). PCR reactions consisted of 10X PCR buffer, 0.4 M_M dNTP, 1.5 M_M $MgCl_2$, 4 μg BSA, 0.25 μ_M each of forward and reverse primers, 0.5–1 μl of template DNA, and 0.5 μl *Taq* DNA polymerase for a final volume of 10 μl. The primers for microsatellite loci were: *CAG133*, F-5′-ATT TTT GCG TCA ACG TGA-3′ and R-5′-ATG AGG ATG TTG TCC AGT TT-3′; *GmcCAG2*, F-5′-GCT TTT CTC GTC CCA TAA-3′ and R-5′-GCG TTG TTG ATGACT GTG-3′; *Gmm22*, F-5′-CGT AAA CGC GGG CTT GT-3′ and R-5′-CAA TTT GGC TGG CTG TCC-3′; *Gmm5B*, F-5′-GAA TTG TTA TGA GTG CAT GT-3′ and R-5′-ATG CGA CAC GAC ACA ATA AG-3′; *Gmm9B*, F-5′-TTT CCT ATA TTG CGA TTA-3′ and R-5′-CGT TTA CGT TAC CCA GAA-3′; *GmsCAG2*, F-5′-GCT TTT CTC GTC CCA TAA-3′ and R-5′-GCG TTG TTG ATG ACT GTG-3′ (Baker & Krafsur, 2001). The PCR reactions were the same as for mitochondrial amplifications except we used 0.25 University of Biolase® polymerase, no BSA, and 0.5 μ_M of each primer for each 10 μl reaction. Amplifications were performed in a PTC-100 programmable thermal cycler (MJ Research, Waltham MA).

Acrylamide gel electrophoresis and silver staining of DNA

We used the methods of Black & DuTeau (1996) in electrophoresis and staining. For the SSCP procedure, after the PCR reactions were completed, 6 μl of loading buffer (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, and 0.05% xylene cyanol) was added to the reaction mixture. The mixtures were then heated in a thermocycler to 95 °C for 4 min, cooled to 0 °C, and immediately loaded on a native polyacrylamide gel. The electrophoretic separations were done by using Hoefer vertical slab gels which were 9% acrylamide in a 1 : 19 bi-acrylamide : acrylamide ratio. Gels contained 5% glycerol and 1X TBE. Gels were run at 2 °C at 250 V for 16 h at which time the xylene cyanol migrated about 13 cm. The lane markers were phiX174 DNA/*Hin*fI (Promega G1751). Microsatellite PCR products were electrophoresed on 0.4 mm thick sequencing gels at 50 °C (Bio-Rad, Hercules, CA). Gels were fixed in 10% acetic acid and the DNA stained with silver.

Phenotypes on gels were scored and photographed with a digital camera (BioVideo-500, Bioimaging Technologies, Brookfield, WI) connected to a Macintosh 6300 computer using the NIH Image program (developed at the U.S. National Institutes of Health and available from the Internet at: <http://rsb.info.nih.gov/nih-image/>.

Sequencing

Did putative alleles inferred from SSCP patterns on acrylamide gels correspond to unique nucleotide sequences? DNA from flies showing the same allele at a locus was amplified in 30 μl PCR reactions. Product concentrations of 5 ng per 100 bp were submitted to the Iowa State Sequencing and Synthesis Facility and sequenced using the Sanger automated dideoxy method. Reactions were run using the Applied Biosystems Prism Big Dye terminator cycle sequencing kit with *AmpliTaq DNA polymerase FS*, and electrophoresed on an Applied Biosystems Prism 377 DNA sequencer. Each PCR fragment was sequenced in both directions using the SSCP primers, chromatograms were assembled, and conflicting base calls were corrected by the program Auto Assembler 1.4.0. Five presumptive alleles at *12S* and four at *16S2* were sequenced. Variation among sequences was identified using SeqApp 1.9 (Gilbert, 1992). No gaps were detected in the sequences. Two presumptive alleles at *16S2* were found to have identical sequences and the haplotype frequencies were adjusted accordingly. Of approximately 721 base pairs sequenced over the two loci, twenty-two (3.1%) were variable. We have sequenced a sum of forty gel phenotypes (presumptive alleles) from *morsitans* group flies and found four of these gel phenotypes to have homologous nucleotide sequences. Each of five putative alleles were found to have two different sequences and one putative allele showed, on sequencing three flies, three different sequences. Thus our SSCP phenotypes underestimate nucleotide diversity. In the final analysis, we scored eleven alleles in seven combinations, or haplotypes.

Data analysis

Chi-squared contingency tests were used to test for homogeneity of allele frequencies among populations. The unbiased estimate of haplotype diversity is given by $h_e = n(1 - \sum x_k^2)/(n - 1)$, where x is the frequency of haplotype k and n is the number of individuals (Nei, 1987). This is the probability that two randomly chosen individuals in the sampled population have different alleles at that locus. The estimate for microsatellite diversity is, $h_e = 2n(1 - \sum x_k^2)/(2n - 1)$.

Gene flow was estimated in terms of Wright's (1978) fixation index F_{ST}. F_{ST} is the standardized variance in allele frequencies among populations, the actual variance divided by the total limiting variance. An estimate of F_{ST} was provided by using the frequency of rare and private alleles by the method of Slatkin (1985). GENEPOP 3.1c software was used to do the calculation (Raymond & Rousset, 1995). An analogous measure of F_{ST} was provided by

θ (the 'coancestry coefficient' of Weir & Cockerham (1984). θ, like F_{ST}, is the correlation of two randomly chosen haplotypes in subpopulations relative to the total population. θ was estimated at three hierarchical levels, populations within regions, regions within the total, and populations relative to the total. Analysis of variance on haplotype and genotypic frequencies according to the methods of Weir (1996) and Excoffier *et al*. (1992) was performed on the combined dataset by using *Arlequin* (Schneider *et al.*, 1997). Pair-wise θ was also calculated using *Arlequin*. In this paper, θ is identified as F_{ST} because they measure the same thing.

The foregoing models allow estimates of the average amount of gene flow among populations by using Wright's (1951) island model of population structure. The mean number of migrant flies exchanged among populations per generation *Nm* can be obtained from the relationship, $F_{ST} \approx (4Nm + 1)^{-1}$ rearrangement of which leads to $Nm \approx (1 - F_{ST})/4F_{ST}$. For mitochondrial haplotypes the appropriate expression is $(1 - F_{ST})/2F_{ST}$ and it represents only female migrants.

We used the software *Bottleneck* ver. 1.2 (Piry *et al.*, 1999) to test hypotheses that populations were in mutation-drift equilibrium at selectively neutral loci (Cornuet & Luikart, 1996). Departures from equilibrium can be caused by recent bottlenecks in population size or recent increases in population size (Maruyama & Fuerst, 1985).

Acknowledgements

Research supported in part by USPHS grant no. AI40048. This is journal paper no. 18409 of the Iowa Agricultural and Home Economics Experiment Station, Project 3447 and supported in part by Hatch Act and State of Iowa funds. Many thanks to Sarah Byers for laboratory assistance, Mike Baker for assembling and processing chromatograms, and W. C. Black IV for the Mantel FORTRAN program.

References

Avise, J.C. (1994) *Molecular Markers, Natural History and Evolution* London: Chapman & Hall.

Baker, M.D. and Krafsur, E.S. (2001) Identification and properties of microsatellite markers in the tsetse fly *Glossina morsitans sensu lato* (Diptera: Glossinidae). *Mol Ecol Notes*, in press.

Black, W.C. and DuTeau, N.M. (1996) RAPD-PCR and SSCP analysis for insect population genetic studies. In: *Molecular Biology of Insect Disease Vectors: a Methods Manual* (Crampton, J.M., Beard, C.B. and Louis, C., eds), pp. 361–383. Chapman & Hall, London.

Cornuet JM, Luikart G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 1996;144:2001–2014. [PubMed: 8978083]

Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 1992;131:479–491. [PubMed: 1644282]

Ford, J. (1971) *The Role of Trypanosome in African Ecology. A Study of the Tsetse Fly Problem* George Allen & Unwin, London.

Gaggiotti OE, Lange O, Rassmann K, Gliddons C. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. Mol Ecol 1999;8:1513–1520. [PubMed: 10564457]

Gilbert, D.G. (1992) *SeqApp, a Biological Sequence Editor and Analysis Program for Macintosh Computers* Published electronically on the Internet, available at: <http://iubio.bio.indiana.edu/soft/molbio/seqapp/>.

Gooding RH. Genetic variation in tsetse flies and implications for trypanosomiasis. Parasitol Today 1992;8:92–95. [PubMed: 15463582]

Hale LR, Singh RS. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. IV. Mitochondrial DNA variation and the role of history vs. selection in the genetic structure of geographic populations. Genetics 1991;129:103–117. [PubMed: 1682210]

Jarne P, Lagoda JL. Microsatellites, from molecules to populations and back. Trends Ecol Evol 1996;11:424–429.

Jordan, A.M. (1993) Tsetse-flies (Glossinidae). In: *Medical Insects and Arachnids* (Lane, R.P. and Crosskey, R.W., eds), pp. 333–382. Chapman & Hall, New York.

Krafsur ES, Griffiths N. Genetic variation at structural loci in the *Glossina morsitans* species group. Biochem Genet 1997;35:1–11. [PubMed: 9238514]

Krafsur ES, Griffiths N, Brockhouse CL, Brady J. Breeding structure of *G. pallidipes* (Diptera: Glossinidae) populations in East and southern Africa. Bull Entomol Res 1997;87:67–73.

Krafsur ES, Madsen M, Wohlford DL, Mihok S, Griffiths NT. Population genetics of *G. morsitans submorsitans.* Bull Entomol Res 2000;90:329–335. [PubMed: 11020791]

Krafsur ES, Wohlford DL. Breeding structure of *Glossina pallidipes* populations evaluated by mitochondrial variation. J Hered 1999;90:635–642. [PubMed: 10589514]

Maruyama T, Fuerst PA. Population bottlenecks and non-equilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. Genetics 1985;111:675–689. [PubMed: 4054612]

Moloo SK, Kabata JM, Waweru F, Gooding RH. Selection of susceptible and refractory lines of *Glossina morsitans centralis* for *Trypanosoma congolense* infection and their susceptibility to different pathogenic *Trypanosoma* species. Med Vet Entomol 1998;12:391–398. [PubMed: 9824823]

Nei, M. (1987) *Molecular Evolutionary Genetics* Columbia University Press, New York.

Noor MAF, Schug MD, Aquadro CF. Microsatellite variation in populations of *Drosophila pseudoobscura* and *Drosophila persimilis.* Genet Res Camb 2000;75:25–35.

Piry S, Luikart G, Cornuet JM. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered 1999;90:502–503.

Raymond M, Rousset F. GENEPOP (Version 3.1c): population genetics software for exact tests and ecumenicism. J Hered 1995;86:248–249.

Rogers D. Study of a natural population of *Glossina fuscipes fuscipes* Newstead and a model for fly movement. J Anim Ecol 1977;46:309–330.

Rogers DJ, Randolph SE. Population ecology of tsetse. Annu Rev Entomol 1985;30:197–216. [PubMed: 3882049]

Schneider, S., Kueffer, J.-M., Roessli, D. and Excoffier, L. (1997) *Arlequin ver. 1.1: A Software for Population Genetic Data Analysis* Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schug MD, Mackay TFC, Aquadro CF. Low mutation rates of microsatellite loci in *Drosophila melanogaster.* Nature Genet 1997;15:99–102. [PubMed: 8988178]

Shahjahan RM, Hughes KJ, Leopold RA, DeVault JD. Lower incubation temperature increases yield of insect genomic DNA isolated by the CTAB method. Biotechniques 1995;19:333–334.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequence and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am 1994;87:651–701.

Slatkin M. Rare alleles as indicators of gene flow. Evolution 1985;39:53–65.

Sokal, R.R. and Rohlf, F.J. (1995) *Biometry: the Principles and Practice of Statistics in Biological Research*, 3rd edn. W.H. Freeman, New York.

Solano P, de La Rocque S, de Meeus T, Cuny G, Duvallet G, Cuisance D. 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 2000;9:433–439. [PubMed: 10971721]

Swofford DL, Selander RB. BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J Hered 1981;72:281–283.

Torr SJ, Parker AG, Leigh-Brown G. The responses of *Glossina pallidipes* Austen (Diptera: Glossinidae) to odour-baited traps and targets in Somalia. Bull Entomol Res 1989;79:99–108.

Vale GA, Hursey BS, Hargrove JW, Torr SJ, Allsopp R. The use of small plots to study populations of tsetse (Diptera: Glossinidae). Insect Sci Applic 1984;5:403–410.

Vale GA, Lovemore DF, Flint S, Cockbill GF. Odour baited targets to control tsetse flies, *Glossina* spp. (Diptera: Glossinidae) in Zimbabwe. Bull Entomol Res 1988;78:31–49.

Weir, B.S. (1996) *Genetic Data Analysis II* Sinauer Assoc., Sunderland, MA.

Weir BS, Cockerham CC. Estimating *F*-statistics for the analysis of population structure. Evolution 1984;38:1358–1370.

Wohlford DL, Krafsur ES, Griffiths NT, Marquez JG, Baker MD. Genetic differentiation of some *Glossina morsitans morsitans* populations. Med Vet Entomol 1999;13:377–385. [PubMed: 10608226]

Wright S. The genetical structure of populations. Ann Eugen 1951;15:323–354.

Wright, S. (1978) Variability within and among natural populations. In: *Evolution and the Genetics of Populations*, Vol. 4. University of Chicago Press, Chicago.

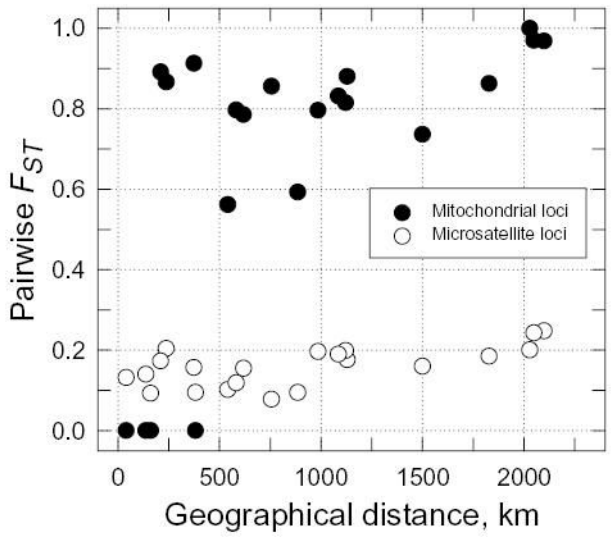

Figure 1.
Pair-wise mitochondrial and microsatelllite genetic distance measures plotted on the geographical distances between them.

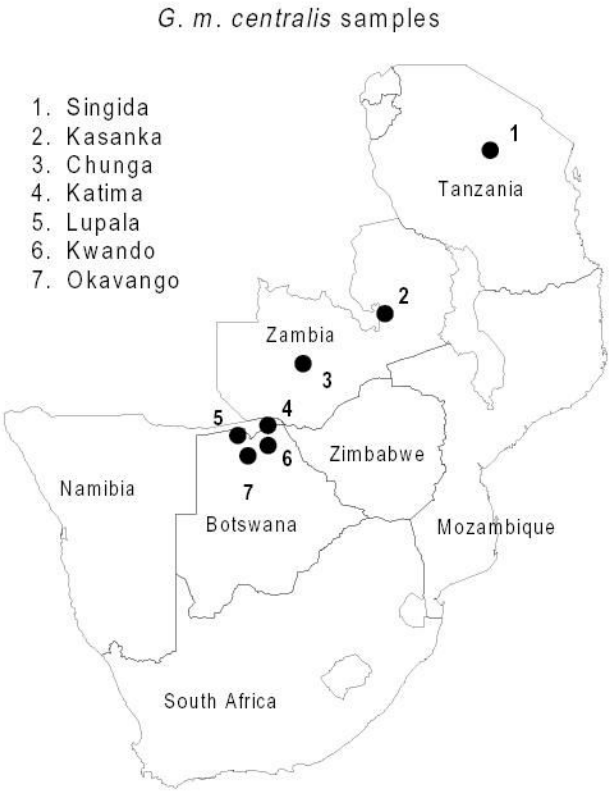

Figure 2.
Southern Africa showing approximate locations of *G. m. centralis* sampling sites. 1, Singida; 2, Kasanka; 3, Chunga; 4, Katima; 5, Lupala; 6, Kwando; 7, Okavango.

Table 1

Haplotype diversity in *G. m. centralis*.

Haplotype	Singida	Kasanka	Chunga	Katima	Lupala	Kwando	Okavango	Total
1	24	0	0	0	0	0	0	24
2	0	0	0	0	35	41	98	174
3	0	0	0	0	0	1	2	3
4	0	0	21	11	0	0	0	32
5	0	0	9	3	0	0	0	12
6	0	14	0	0	0	0	0	14
7	0	6	0	0	0	0	0	6
Total	24	20	30	14	35	42	100	265
Diversity	0	0.442	0.434	0.363	0	0.048	0.040	0.543
SD	—	0.087	0.070	0.130	—	0.045	0.027	0.156

Appendix

Microsatellite allelic frequencies in *G. m. centralis*.

	Population						
Locus	**Singida**	**Kasanka**	**Chunga**	**Katima**	**Lupala**	**Kwando**	**Okavango**
CAG133							
N	21	15	23	8	15	21	24
A	0.024	0.133	0.000	0.000	0.100	0.000	0.229
B	0.262	0.267	0.326	0.375	0.500	0.500	0.521
C	0.048	0.133	0.261	0.188	0.100	0.357	0.229
D	0.333	0.233	0.261	0.375	0.233	0.071	0.021
E	0.333	0.233	0.065	0.063	0.067	0.024	0.000
F	0.000	0.000	0.087	0.000	0.000	0.048	0.000
GmcCAG							
N	23	21	24	22	24	19	22
A	0.087	0.000	0.000	0.023	0.146	0.053	0.023
B	0.913	0.952	0.833	0.727	0.521	0.763	0.864
C	0.000	0.000	0.063	0.182	0.313	0.132	0.000
D	0.000	0.048	0.021	0.068	0.021	0.026	0.114
E	0.000	0.000	0.083	0.000	0.000	0.026	0.000
Gmm22							
N	23	15	21	18	19	21	12
A	0.022	0.033	0.048	0.028	0.000	0.000	0.125
B	0.000	0.133	0.000	0.056	0.026	0.071	0.000
C	0.000	0.600	0.214	0.028	0.026	0.071	0.208
D	0.000	0.200	0.310	0.194	0.316	0.119	0.667
E	0.000	0.000	0.071	0.139	0.105	0.214	0.000
F	0.370	0.000	0.024	0.028	0.342	0.214	0.000
*G	0.000	0.000	0.119	0.000	0.000	0.000	0.000
H	0.196	0.000	0.071	0.000	0.026	0.000	0.000
I	0.000	0.000	0.071	0.000	0.000	0.000	0.000
J	0.000	0.033	0.024	0.056	0.000	0.000	0.000
K	0.413	0.000	0.000	0.028	0.053	0.214	0.000
L	0.000	0.000	0.000	0.444	0.105	0.048	0.000
Gmm5B							
N	24	21	25	20	22	21	23
A	0.000	0.000	0.080	0.000	0.000	0.048	0.043
B	0.000	0.000	0.020	0.000	0.000	0.071	0.109
C	0.500	0.286	0.120	0.400	0.250	0.143	0.022
D	0.000	0.381	0.080	0.325	0.045	0.214	0.239
E	0.000	0.000	0.360	0.000	0.000	0.095	0.022
F	0.000	0.000	0.100	0.000	0.386	0.286	0.370
G	0.000	0.024	0.080	0.050	0.159	0.095	0.043
H	0.458	0.000	0.080	0.000	0.023	0.000	0.000
I	0.042	0.143	0.000	0.000	0.136	0.000	0.130
J	0.000	0.167	0.000	0.225	0.000	0.000	0.022
Gmm9B							
N	17	18	21	14	17	21	24
A	0.118	0.000	0.024	0.393	0.000	0.000	0.000
B	0.000	0.306	0.048	0.000	0.000	0.000	0.000
C	0.000	0.000	0.095	0.000	0.059	0.000	0.063
D	0.000	0.056	0.048	0.000	0.441	0.000	0.250
E	0.000	0.000	0.071	0.000	0.000	0.286	0.000
F	0.618	0.000	0.095	0.000	0.000	0.000	0.000
G	0.000	0.000	0.071	0.071	0.029	0.000	0.021
H	0.000	0.083	0.024	0.000	0.000	0.310	0.000
I	0.000	0.028	0.048	0.000	0.441	0.024	0.125
J	0.000	0.000	0.167	0.000	0.029	0.119	0.542
K	0.000	0.000	0.167	0.071	0.000	0.071	0.000
L	0.000	0.000	0.095	0.000	0.000	0.119	0.000
M	0.265	0.528	0.048	0.464	0.000	0.024	0.000
N	0.000	0.000	0.000	0.000	0.000	0.048	0.000
GmsCAG							
N	21	12	24	20	10	21	19
A	0.095	0.042	0.083	0.025	0.200	0.000	0.026
B	0.810	0.958	0.833	0.900	0.450	0.071	0.763
C	0.048	0.000	0.063	0.075	0.000	0.619	0.158
D	0.048	0.000	0.021	0.000	0.350	0.071	0.053
E	0.000	0.000	0.000	0.000	0.000	0.214	0.000
F	0.000	0.000	0.000	0.000	0.000	0.024	0.000

Table 2

Pair-wise genetic distance measures in terms F_{ST} of in the upper diagonal and equivalent dispersal rates* in the lower diagonal.

	Singida	Kasanka	Chunga	Katima	Lupala	Kwando	Okavango
Singida	—	0.796	0.763	0.863	1	0.97	0.968
Kasanka	0.128	—	0.562	0.593	0.832	0.815	0.88
Chunga	0.155	0.389	—	0	0.797	0.785	0.856
Katima	0.079	0.343	>> 10	—	0.892	0.867	0.913
Lupala	0	0.101	0.128	0.061	—	0	0
Kwando	0.016	0.113	0.137	0.077	>> 10	—	0
Okavango	0.016	0.068	0.084	0.048	>> 10	>> 10	—

*Average number of migrants ~$(1 - F_{ST})/2F_{ST}$.

Table 3

Analysis of variance in *G. m. centralis* mitochondrial haplotype frequencies by the method of Excoffier *et al.* (1992). All haplotypes are considered equidistant.

Source	d.f.	Variance	% of total	*F* statistics	Significance*
Among groups†	2	0.3404	74.5	F_{SG} = 0.475	<0.007
Among populations within groups	4	0.0553	12.1	F_{GT} = 0.745	<0.0001
Within populations	258	0.0612	13.4	F_{ST} = 0.866	<0.0001
Total	264	0.4459			

*Probability of obtaining a greater variance and *F* by chance.

†Groups are (1) Tanzania; (2) Kasanka, Chunga, Katima; (3) Lupala, Kwando, and Okavango.

Table 4

Diversities averaged over six microsatellite loci in *G. m. centralis*.

Population	Mean sample size per Locus	Mean no. of alleles per locus	Mean heterozygosity	
			Observed	Expected
Singida	21.5	3.5 ±0.4	0.549 ± 0.096	0.499 ± 0.086
Kasanka	17.0	4.0 ± 0.6	0.552 ± 0.155	0.494 ± 0.132
Chunga	23.0	7.2 ± 1.4	0.719 ± 0.122	0.661 ± 0.116
Katima	17.0	4.7 ± 0.9	0.621 ± 0.114	0.576 ± 0.090
Lupala	17.8	5.2 ± 0.7	0.823 ± 0.071	0.691 ± 0.027
Kwando	20.7	6.2 ± 0.5	0.761 ± 0.081	0.686 ± 0.074
Okavango	20.7	4.7 ± 0.7	0.603 ± 0.093	0.538 ± 0.079

Table 5

Microsatellite loci: diversities and *F* statistics by the method of Weir and Cockerham.

Locus	No. alleles	Diversity	F_{IS}	F_{ST}	F_{IT}
GpCAG133	6	0.745	−0.178	0.073	−0.092
GmcCAG	5	0.352	−0.167	0.099	−0.052
Gmm22	12	0.858	−0.093	0.184	0.109
Gmm5B	10	0.855	−0.170	0.159	0.016
Gmm9B	14	0.898	0.022	0.257	0.273
GmcCAG	6	0.502	−0.196	0.322	0.190
Mean	8.83	0.702	−0.120	0.185	0.087
	Jackknife estimates over loci				
	Mean	−0.119	0.186	0.090	
	SD	0.038	0.035	0.063	

Table 6

Pair-wise genetic distance measures at microsatellite loci in terms of F_{ST} in the upper diagonal and equivalent dispersal rates* in the lower diagonal.

	Singida	Kasanka	Chunga	Katima	Lupala	Kwando	Okavango
Singida	—	0.197	0.16	0.185	0.201	0.243	0.248
Kasanka	1.02	—	0.102	0.095	0.19	0.199	0.176
Chunga	1.31	2.21	—	0.095	0.119	0.155	0.078
Katima	1.1	2.39	2.39	—	0.173	0.204	0.157
Lupala	0.99	1.06	1.85	1.19	—	0.132	0.093
Kwando	0.78	1.01	1.36	0.98	1.64	—	0.14
Okavango	0.76	1.17	2.94	1.34	2.44	1.53	—

* Average number of migrants ~$(1 - F_{ST})/4F_{ST}$.

Table 7

Analysis of variance in *G. m. centralis* microsatellite allelic frequencies by the method of Excoffier *et al*. (1992). All alleles are considered equidistant.

Source of variation	d.f.	Variance	Percentage of total	*F* statistics	Significance
Among groups*	2	0.1272	7.2	$F_{SG} = 0.114$	$P < 0.0001$
Among populations within groups	4	0.1186	10.6	$F_{GT} = 0.071$	$P < 0.015$
Within populations	303	1.4637	82.3	$F_{ST} = 0.177$	$P < 0.0001$
Total	309	1.7614			

*Groups are: (1) Tanzania; (2) Kasanka, Chunga, Katima; (3) Lupala, Kwando, and Okavango.