

Original article

A comparison of *Anopheles gambiae* and *Plasmodium falciparum* genetic structure over space and time

Franck Prugnolle^{a,*}, Patrick Durand^a, Koella Jacob^c, Fabien Razakandrainibe^b,
Céline Arnathau^a, Diana Villarreal^a, François Rousset^d, Thierry de Meeûs^a, François Renaud^a

^a *Génétique et Evolution des Maladies Infectieuses, UMR 2274 IRD-CNRS-UMI, Centre IRD, 911 avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France*

^b *Laboratoire de Parasitologie Evolutive, Unité Mixte de Recherche-Centre National de la Recherche Scientifique 7103, UPMC, 75252 Paris, France*

^c *Division of Biology, Imperial College London, Silwood Park, Ascot, Berkshire SL5 7PY, UK*

^d *Institut des Sciences de l'Evolution Unité Mixte de Recherche, Centre National de la Recherche Scientifique 5554, 34095 Montpellier Cedex 5, France*

Received 27 September 2007; accepted 3 December 2007

Available online 12 January 2008

Abstract

Population genetic structure and subdivision are key factors affecting the evolution of organisms. In this study, we analysed and compared the population genetic structure of the malaria parasite *Plasmodium falciparum* and its mosquito vector *Anopheles gambiae* over space and time in the Nianza Province, near Victoria Lake in Kenya. The parasites were collected from mosquitoes caught in six villages separated by up to 68 km in 2002 and 2003. A total of 545 oocysts were dissected from 122 infected mosquitoes and genotyped at seven microsatellite markers. Five hundred and forty-seven mosquitoes, both infected and uninfected, were genotyped at eight microsatellites. For the parasite and the vector, the analysis revealed no (or very little) genetic differentiation among villages. This may be explained by high local population sizes for the parasite and the mosquito. The small level of genetic differentiation observed between populations may explain the speed at which antimalarial drug resistance and insecticide resistance spread into the African continent.

© 2008 Elsevier Masson SAS. All rights reserved.

Keywords: *Plasmodium falciparum*; *Anopheles gambiae*; Population genetics; Africa

1. Introduction

Population genetic structure and subdivision have long been recognised as important factors affecting the evolution of organisms [1]. Nevertheless, our knowledge of the genetic structure of hosts and parasites remains particularly limited in natural conditions even for systems of medical or veterinary importance [2].

Plasmodium falciparum, the agent of the most malignant form of human malaria, is not an exception. While this pathogen is one of the most deadly, with more than 2 million people killed each year especially in Africa, we know little about

the parasite population structure as well as that of its vector hosts in the same areas [3,4]. The sizes of their populations (which largely determines the level of genetic drift and thus the efficiency of selection) and their connectivity (which determines the speed at which a favourable mutation may spread through different areas) are among the basic population features that remain largely unknown but are fundamental to their evolution, co-evolution [2] and thus responses to changes in their environments (e.g. [5]). Such knowledge is particularly important in the context of malaria control, as it would help to understand and predict, for example, the evolution of drug resistance in the parasite [6] and the spread of transgenic resistance genes in the vector [7].

In this study, we analysed the population genetic structure of *P. falciparum* and its vector *An. gambiae* in six villages in an area of Kenya with intense and permanent transmission, over two years of sampling.

* Corresponding author. Tel.: +33 (0)4 67 41 61 80; fax: +33 (0)4 67 41 62 99.

E-mail address: prugnoll@mpl.ird.fr (F. Prugnolle).

2. Material and methods

2.1. Sampling sites

A. gambiae and *P. falciparum* were collected in the Nyanza province in Kenya. Six villages (Kapsulu, Koderia, Rangwe, Ringa, Rota and Sondu) near Victoria Lake were studied. All villages were separated by less than 70 km. For more details on the transmission sites, refer to [8]. In all villages, malaria transmission is intense all the year along.

2.2. Mosquito collection

Resting, blood-fed *An. gambiae* females were collected with aspirators from the walls of houses early in the morning during the main rainy seasons of 2002 and 2003, between mid-March and June. Mosquitoes were collected each year for ten weeks in each of the six villages. Mosquitoes were brought to an insectary (maintained at ambient temperature, humidity and light), fed with a sugar solution *ad libitum*, and maintained for 7 days in small pots.

2.3. DNA mosquitoes and *P. falciparum* oocyst extractions

On the 7th day after capture, we dissected the mosquito midguts and checked them with a microscope for the presence of oocysts. Infected midguts were stored in absolute ethyl alcohol and each oocyst was isolated under a Leica DMIRB microscope. Dissected oocysts were individually preserved at -20 °C in absolute ethyl alcohol until their DNA extraction using the DNeasy Tissue Kit following the manufacturer's instructions (Qiagen, CA). Uninfected and infected female mosquitoes (after gut removal) were stored in absolute ethyl alcohol. Genomic DNA was extracted from individual female mosquitoes and the rDNA-PCR method was used to determine the species identity for specimens within the *An. gambiae* species complex [9].

2.4. Microsatellite loci and genotype scoring

We investigated 8 and 7 microsatellite loci for *An. gambiae* and *P. falciparum*, respectively (Table 1). PCR amplifications followed the protocols of [10] and [11] for the mosquitoes and [12] for the parasites. Fluorescence-labelled PCR products were sized on ABI Prism310 genetic analyzer (Applied Biosystems), with a Genescan500LIZ internal size standard.

2.5. Statistical analyses

2.5.1. Population genetic structure

Genetic analyses were performed on (i) 547 mosquitoes collected in 2002 and 2003 in six villages (Table 2) and (ii) 545 *P. falciparum* oocysts dissected from 122 infected *An. gambiae* collected from the same villages but one (Sondu). In this latter village too few infected mosquitoes were collected to allow the genetic analyses.

Table 1

Microsatellite loci characteristics for *An. gambiae* (loci coded from A to H) and for *P. falciparum* (loci coded from 1 to 7)

Locus	Code	Cytologic location	Repeat motif	GenBank acc. no.	Ref.
AgXH1D1	A	X:1d	(CCA)	Z72021	[10]
Ag2H22C1	B	II:22c	(TG)	NA	[10]
Ag2H147	C	II:2R19	(GT)	Z72052	[10]
Ag2H46	D	II:7a	(GT)	Z72020	[11]
AgX2A1	E	X:2a	(GT)	NA	[10]
Ag2H143	F	II:2La	(TC)	Z72050	[10]
Ag3H29C1	G	III:29c	(TGA)	NA	[10]
Ag3H119	H	III:35b	(GT)	Z72043	[10]
POLYa	1	Chr. 4	(TAA)	G37809	[12]
TA60	2	Chr. 13	(TAA)	G38876	[12]
ARA2	3	Chr. 11	(TAA)	G37848	[12]
Pfg377	4	Chr. 12	(TAA)	G37851	[12]
PfPK2	5	Chr. 12	(ATT)	G37852	[12]
TA87	6	Chr. 6	(CAA) (TAA)	G38838	[12]
TA109	7	Chr. 6	(ACT)	G38842	[12]

NA, not available.

FSTAT V.2.9.3 (updated from [13]) was used to estimate Nei's [14] unbiased estimate of genetic diversities (H_s) and to test for deviations from Hardy–Weinberg expectations within mosquito and parasite populations. Deviations from Hardy–Weinberg proportions were tested using f , the Weir and Cockerham's [15] estimator of Wright F_{IS} [1] as a statistic and 10,000 within-samples permutations of alleles among individuals to compute the P -value.

Genetic differentiation among villages was estimated with different methods according to the species concerned. For mosquitoes, differentiation among all six populations in each time interval and between pairs of populations within and between temporal periods, was estimated using Weir and Cockerham's estimator θ [15] of Wright's F_{ST} [1] calculated by the program FSTAT V.2.9.3. Population differentiation was tested using the genotypic log-likelihood ratio G test described in Goudet et al. [16] and by performing

Table 2

Summary statistics for the six studied populations of *An. gambiae*

Populations	Year	N	H_o	H_s	F_{IS}
Kapsulu (KAPS)	2002	5	0.475	0.633	0.273*
	2003	55	0.627	0.693	0.095***
Koderia (KOD)	2002	88	0.601	0.704	0.147***
	2003	38	0.604	0.730	0.175***
Rangwe (RAN)	2002	16	0.515	0.653	0.216***
	2003	47	0.621	0.699	0.113***
Ringa (RIN)	2002	40	0.574	0.661	0.132***
	2003	56	0.600	0.704	0.149***
Rota (ROT)	2002	64	0.616	0.692	0.111***
	2003	89	0.607	0.701	0.134***
Sondu (SON)	2002	12	0.541	0.653	0.177**
	2003	37	0.565	0.724	0.223***

N , number of mosquitoes genotyped; H_o , observed heterozygosity; H_s , Nei's [13] estimate of local genetic diversity; F_{IS} , heterozygosity deviation from Hardy–Weinberg expectations. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

10,000 among-samples permutations of individual genotypes. For the parasite, another procedure was used to take into account for the different levels of hierarchy (e.g. [17]) occurring in our samples. The first level of subdivision is the oocyst sub-population carried by an individual mosquito (the infrapopulation [8]); the second is the site (or village) where infected mosquitoes were collected. Population differentiation over all sites as well as between pairs of sites within and between temporal periods was estimated using HIERFSTAT version 0.03-2, a package of the statistical software R (R Development Core Team, 2004) that computes hierarchical F -statistics from any number of hierarchical levels [18]. The significance of genetic differentiation among sites or among temporal samples (hereafter called F_{VT}) was tested by randomly permuting 1000 times the oocyst sub-populations among the different samples of interest (the sites or the temporal samples).

For mosquitoes and parasites, an isolation-by-distance model of population genetic structure was tested by regressing $\theta/(1 - \theta)$ (or $F_{VT}/(1 - F_{VT})$) against the natural logarithm of geographical distance between population pairs [19]. Distances between locations were calculated with the Encarta program using the geo-localisation of each site. The statistical significance of the regression was tested with a Mantel test with 5000 permutations using FSTAT V. 2.9.3.

Genotypic disequilibria among pairs of loci were tested with 10,000 genotype permutations. The significance of the association between each pair of loci over all samples was determined with the log-likelihood ratio G statistic.

For all multiple test procedures (e.g. pairwise F_{ST} and linkage disequilibrium), the significance levels were adjusted with the Bonferroni correction [20].

2.5.2. Population genetic structure of the “infected” and “uninfected” mosquitoes

We investigated possible genetic differences among infected and uninfected mosquitoes within each population using a particular permutation procedure available in FSTAT V. 2.9.3 (updated from [13]; “Sex-Biased Disp” procedure adapted for “infected” and “uninfected” hosts). Its principle is the following: if “infected” and “uninfected” individuals harbour the same structure, the statistics used to describe it (e.g. F_{ST} , H_s etc.) should not differ between them. Let X_m and X_f be the statistic of interest for “infected” and “uninfected” mosquitoes, respectively. Then, the test proceeds as follows. The statistic of interest (F_{IS} , F_{ST} , H_o (the observed heterozygosity), H_s) for both groups (infection status) and the absolute value of their difference $\Delta X_{obs} = |X_m - X_f|$ are calculated over all populations. The infection status is then randomly assigned 10,000 times to each individual (keeping the individuals in their original population and the infected/uninfected ratio in each sample constant). The P -value of the test is then computed as the proportion of randomly obtained values of ΔX that are equal to or higher than the observed value. For maximal power, the analyses were performed over all villages for each year.

3. Results

3.1. *Anopheles gambiae*

According to PCR identifications, all 547 mosquitoes analysed belonged to the S molecular form of *An. gambiae sensu stricto*. Over all samples and years, the number of alleles of the 8 microsatellites ranged from 4 to 21 and the genetic diversity H_s ranged from 0.65 to 0.73 within villages. The genetic diversity was slightly higher in 2003 than in 2002 (Table 2; Wilcoxon signed rank test: $P = 0.031$). No linkage disequilibrium was found between any pair of loci within mosquito populations after Bonferroni corrections (Table 4).

Strong and significant departures from Hardy-Weinberg expectations within each village were observed at each time period (Table 2). Overall, F_{IS} (f) within villages ranged from 0.095 to 0.27 indicating large deficits in heterozygotes. Fig. 1 shows F_{IS} values for the eight loci in anopheline populations within each time period. Locus D (Ag2H46) showed high heterozygote deficits in contrast to other loci, which showed less deviation from Hardy-Weinberg expectations. Null alleles have been found at this locus in the literature [21]. When we excluded this locus from the analysis, the mean F_{IS} (f) over all loci decreased, but remained highly significant (Fig. 1).

Only a few pairs of populations showed significant differentiation ($\theta > 0$) within each time period (see Table 5) after Bonferroni corrections. Overall, θ among populations was 0.006 for the 2002 samples and 0.002 for the 2003 samples. Only the 2002 θ -value was significantly different from 0. No significant patterns of isolation by distance were detected at the scale under scrutiny for either period of sampling (Mantel test: 2002: $P = 0.53$; 2003: $P = 0.082$). Within each village, no genetic differentiation was observed between the two periods of sampling (Table 5).

Finally, over all villages and years, we observed no significant difference in the distribution of genetic variability between infected (In) and uninfected (Un) mosquitoes (2002: for the F_{IS} : In = 0.135; Un = 0.139; $P = 0.93$; for the F_{ST} : In = 0.017; Un = 0.004; $P = 0.20$; for H_o : In = 0.592; Un = 0.595; $P = 0.90$; for H_s : In = 0.685; Un = 0.690; $P = 0.69$; 2003: for the F_{IS} : In = 0.097; Un = 0.153;

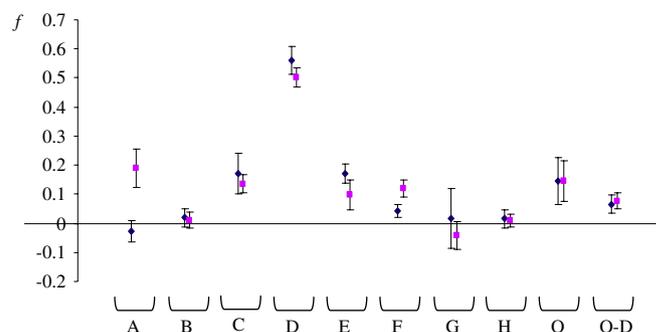


Fig. 1. *An. gambiae* F_{IS} estimates (f) per locus and year of sampling (2002: blue diamonds; 2003: pink squares). A–H refer to the different loci (see Table 1 for the correspondence between the code and the name of the locus). O, over all loci. O-D, over all loci but locus D.

$P = 0.08$; for the F_{ST} : $In = 0.014$; $Un = 0.001$; $P = 0.11$; for H_0 : $In = 0.623$; $Un = 0.602$; $P = 0.33$; for H_s : $In = 0.689$; $Un = 0.711$; $P = 0.08$).

3.2. Plasmodium falciparum

The microsatellite loci of the 545 oocysts genotyped were highly polymorphic. The number of alleles ranged from 7 (locus Pfg377) to 19 (loci PfpK2 and TA109) and the genetic diversity over all oocyst subpopulations within each village (H_T) ranged from 0.738 to 0.817 (Table 3). Deviations from Hardy–Weinberg expectations, as measured by F_{IS} , were detected within oocyst subpopulations (i.e. within mosquitoes) and ranged from 0.093 to 0.31 as previously observed by [8] (Table 3). In contrast to the vector, all pairs of loci were in linkage disequilibrium in *P. falciparum* (over all sites, 21 tests out of 21 significant at $P = 0.05$ after Bonferroni procedure) (Table 4).

In *Plasmodium*, population structure occurs at different levels. Within sites, oocyst subpopulations constituted one level of population structure. Genetic differentiation among oocyst subpopulations was significant (as previously demonstrated [8]) and ranged from 0.036 to 0.397 over the two periods of sampling. The second level of population structure was the site where infected mosquitoes were collected. Once the lower level of differentiation (the oocyst subpopulation) was taken into account, no significant differentiation was observed among villages. No isolation by distance pattern could be found in either 2002 ($P = 0.52$) or 2003 ($P = 0.24$). Within each village, no genetic differentiation was observed between years (Table 5).

4. Discussion

Comparing over space and time the distribution of genetic variability of hosts and parasites brings information on the functioning of each species, especially with respect to those factors that may influence their evolutionary and co-evolutionary dynamics such as their effective population size or migration rate. In this study, we analysed and compared the population

genetic structure of the vector *An. gambiae* and the parasite *P. falciparum* in an area of Kenya with intense transmission.

4.1. An. gambiae population genetic structure

For *An. gambiae*, we observed significant deficits in heterozygous individuals (i.e. $F_{IS} > 0$). Although the F_{IS} -estimates are biased by the null alleles in one of the loci (locus AG2H46) [21], they remained significantly different from 0 when this locus was omitted from the analyses. Similar deviations were reported in all previous microsatellite studies on *An. gambiae* (e.g. [10,11,22–26]) and were considered to represent null alleles. Our results give some insight into this issue with the analysis of linkage disequilibrium (LD), which measures the departure of the association of alleles at different loci from that expected in a population under random mating. LD is expected if members of different subpopulations (Wahlund effect) or different inbred “lines” are pooled in an analysis, as they have different probabilities to carry certain combinations of alleles. In contrast, LD is not expected if heterozygote deficits are only due to null alleles, because all individuals are equally likely to carry a null allele. Thus, as no LD was observed among any pairs of microsatellite loci, our results on heterozygote deficiency are likely to be due to the presence of null alleles.

An. gambiae populations were only very slightly genetically differentiated among villages at the scale under scrutiny (the maximal distance between two populations was about 68 km). This result is consistent with the majority of previous studies on *An. gambiae* in similar regions of Africa—especially in Kenya [10,11,24,25]—showing that the geographical area associated with a population of *An. gambiae* was generally larger than 50 km in diameter.

Such low genetic differentiation observed between populations is indicative of a large local population size of *An. gambiae* in the area, which is further corroborated by the observation that allele frequencies did not vary significantly between the temporal samples (separated by about 12 generations [24]) and that a high level of genetic variability was maintained in the population. Again, this result is congruent with previous studies

Table 3
Summary statistics for the five populations of *P. falciparum*

Populations		N_{im}	N_p	H_0	H_s	H_T	F_{SV}	F_{IS}
Kapsulu (KAPS)	2002	5	16	0.615	0.837	0.817	0.036*	0.331***
	2003	13	36	0.365	0.609	0.796	0.2***	0.248***
Kodera (KOD)	2002	18	181	0.407	0.515	0.787	0.353***	0.093***
	2003	2	3	—	—	—	—	—
Rangue (RAN)	2002	3	12	0.388	0.596	0.775	0.276***	0.276**
	2003	13	29	0.285	0.422	0.75	0.425	0.165
Ringa (RIN)	2002	6	18	0.394	0.526	0.764	0.369***	0.186*
	2003	11	17	0.539	0.646	0.752	0.196***	0.174*
Rota (ROT)	2002	16	54	0.46	0.565	0.782	0.249***	0.134*
	2003	35	179	0.367	0.433	0.757	0.397***	0.133*

N_{im} , number of infected mosquitoes used in the study; N_p , total number of parasites genotyped within infected mosquitoes; H_0 , oocyst subpopulation observed heterozygosity; H_s , oocyst subpopulation expected heterozygosity. H_T , expected heterozygosity over all oocyst subpopulations, within villages; F_{IS} , deviation from Hardy–Weinberg expectations within oocyst subpopulations. F_{SV} , genetic differentiation among oocyst subpopulations. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 4

Linkage disequilibrium: number of significant combinations of pairs of loci on a total of 28 and 21 combinations in *An. gambiae* and *P. falciparum* populations, respectively

		KAPS	KOD	RAN	RIN	ROT	SON	All sites
<i>An. gambiae</i>	2002	0 (<i>N</i> = 5)	0 (<i>N</i> = 88)	0 (<i>N</i> = 16)	0 (<i>N</i> = 40)	0 (<i>N</i> = 64)	0 (<i>N</i> = 12)	0 (<i>N</i> = 225)
	2003	0 (<i>N</i> = 55)	0 (<i>N</i> = 38)	0 (<i>N</i> = 47)	0 (<i>N</i> = 56)	0 (<i>N</i> = 89)	0 (<i>N</i> = 37)	0 (<i>N</i> = 322)
<i>P. falciparum</i>	2002	0 (<i>N</i> = 16)	21 (<i>N</i> = 181)	2 (<i>N</i> = 12)	2 (<i>N</i> = 18)	3 (<i>N</i> = 54)	–	21 (<i>N</i> = 281)
	2003	6 (<i>N</i> = 36)	– (<i>N</i> = 3)	7 (<i>N</i> = 29)	0 (<i>N</i> = 17)	21 (<i>N</i> = 179)	–	21 (<i>N</i> = 264)

Significance was determined after Bonferroni corrections. *N* indicates the number of individuals genotyped per population.

reporting high estimates of local size of *An. gambiae* populations in nearby regions in Kenya [24]. It may explain why only little genetic differentiation is observed between some high-transmission sites of western and eastern Africa, although they are separated by thousands of kilometres [26].

Changes of allelic frequencies over time within populations are frequently used to infer their effective size. One traditionally used method is the one described by Waples [27]. It was, for instance, applied to estimate the effective size of *An. gambiae* in Kenya [24]. Waples' method [27] assumes that sampling is random, that no subdivision of gene pool exists and that selection, mutation and migration are negligible. Those assumptions are quite strong and very likely to be wrong in natural conditions. This is especially true here for the hypothesis of no immigration as we observed no or very little genetic differentiation between the populations sampled. Because immigration is known to strongly skew the estimate of effective size either upward or downward [28], we thus avoided using such a method and simply limited our analysis to the computation of genetic differentiation between temporal samples.

Comparing genetic variation between infected and uninfected hosts may reveal interesting properties regarding the origin of host susceptibility to the parasite and its evolution (see e.g. [29]). For instance, a lower average observed heterozygosity of the infected hosts may reveal a problem of inbreeding depression lowering their ability to resist parasites. A lower F_{ST} among infected hosts could reveal a phenomenon of host local adaptation to the parasite resulting in an increase susceptibility of immigrants compared to residents. In our study, we did not detect any difference in the distribution of genetic variability between infected and uninfected mosquitoes. Although this result may suggest the absence of the aforesaid phenomena (i.e. lower resistance of more homozygous individuals or of immigrants), we cannot completely ruled out the possibility of a lack of power to detect a difference between infected and uninfected mosquitoes.

4.2. *P. falciparum* population genetic structure

The way genetic variability of *P. falciparum* is distributed within villages, both within and among oocyst subpopulations,

Table 5

An. gambiae and *P. falciparum* genetic differentiation

	KAPS	KOD	RAN	RIN	ROT	SON
<i>An. gambiae</i> populations						
KAPS	0.0213	0.0038	0.0032	0.0087*	0.0035	0.0062
KOD	0.0191	−0.0047	−0.0032	−0.0035	0.0012	0.0001
RAN	0.0117	0.0029	−0.002	−0.0028	0.0029	−0.0027
RIN	0.0472	0.0044	−0.0023	0.001	0.001	0.002
ROT	0.0328	0.0034*	0.0061	0.0065**	−0.0019	0.0029
SON	0.0452	0.0099	0.0048	0.002	0.0036	0.0144
Global θ value for 2002 samples	0.006**					
Global θ value for 2003 samples	0.002					
<i>P. falciparum</i> populations						
KAPS	−0.017	0.0015	0.0088	0.0072	−0.014	NA
KOD	−0.053	0.018	−0.0147	0.03	0.027	NA
RAN	0.053	−0.013	0.0006	0.0066	0.009	NA
RIN	−0.032	0.018	−0.009	−0.019	−0.0142	NA
ROT	−0.034	0.001	0.03	0.0058	−0.0031	NA
SON	NA	NA	NA	NA	NA	NA
Global θ_{VT} value for 2002 samples	−0.0049					
Global θ_{VT} value for 2003 samples	−0.0048					

Above diagonal: θ values between pairs of samples collected in 2003. Below diagonal: θ values between pairs of samples collected in 2002. On diagonal (in bold): θ values between populations from the same village between the two sampling years. In all cases, asterisks designate significance at the Bonferroni level with * $P < 0.05$ and ** $P < 0.01$. For each species, below the matrix are global θ values for both sampling periods. For *An. gambiae*, pairwise F_{ST} (θ) were computed using FSTAT V.2.9.3. For *P. falciparum*, pairwise F_{VT} (θ_{VT}) between villages were computed using the hierarchical method implemented in HIERFSTAT.

has already received attention in Razakandrainibe et al. [8]. Contrary to *An. gambiae*, *P. falciparum* displays strong linkage disequilibrium, is locally highly structured (oocyst infrapopulations are genetically highly differentiated) and shows high F_{IS} within infrapopulations, a likely result of (i) high levels of selfing during fusion of gametes [8,30], or (ii) Wahlund effects resulting from the intake, by the mosquito, of two infected blood meals at sufficiently long time intervals to prevent the fusion of gametes coming from the two blood meals [31].

After taking into account the smallest levels of fragmentation, i.e., between subpopulations of oocysts coming from different mosquitoes, no significant genetic differentiation was observed between villages. Therefore, as for *An. gambiae*, the area covered by the present study (more than 50 km of diameter) hosts hundreds of reproductive units of *P. falciparum*, the heterogeneity of which is fully explained by individual hosts and poorly (if at all) by geography. These results are again consistent with those obtained in previous studies in regions of intense transmission and may thus explain why very little genetic differentiation has been reported even between regions separated by thousands of kilometres [12,30].

As for *An. gambiae*, these findings suggest therefore a very large local population size of the parasite in areas of intense transmission. This idea is here further supported by other results, such as the observed high level of genetic variability maintained in the populations and the fact that no significant variation in allelic frequencies was observed between samples collected one year (and thus about 6 parasite generations) apart. Again, we did not attempt to compute an effective size from the temporal samples, as no temporal-sample based method exists for structured samples with obligate extinction (i.e. the oocyst sub-population constitutes one level of structure within each sample but this entity does not persist over time). Pooling the oocyst sub-populations to form only one sample at each time interval would result into a strong overestimate of the variance in allelic frequencies between temporal samples, because of the genetic differentiation existing among oocyst infrapopulations, and hence an estimate of effective size strongly biased downward using Waples' method [27].

The migration of the parasite between the transmission sites is necessarily host-dependent in *P. falciparum*. It may therefore depend on the mobility of either humans or mosquitoes at the scale of our study. Without any further knowledge of the mobility of each infected host in the sampling region, it cannot be known which host affects the dispersal of the parasite more.

5. Conclusions

As described above, the population genetic structures of *An. gambiae* and *P. falciparum* are similar in certain aspects but different in others. While neither shows much genetic differentiation among Kenyan villages more than 50 km apart, they differ in their population structure within villages. In contrast to the vector, *P. falciparum* populations are far from panmictic, as indicated by the strong departure from Hardy–Weinberg

expectations (high positive F_{IS}), the level of differentiation observed between oocyst infrapopulations and the extent of linkage disequilibria observed in its genome.

This may have several consequences for the evolution of the parasite compared to its mosquito host. While population structure does not affect the probability that a new advantageous mutation is fixed in a population, it can greatly affect its time to fixation, i.e. the speed at which it spreads in the population [32]. In particular, the more structured infrapopulations are, the longer the time to fixation should be, i.e. the more slowly the mutation should spread [32]. Linkage disequilibrium can also affect positive selection if several mutations are selected at the same time in the genome [33]. The efficiency of selection acting simultaneously at several linked sites may thus be considerably reduced because of the high level of linkage disequilibrium occurring in the genome.

The dynamics of new advantageous mutations will thus differ in the population of the mosquito or the parasite. Such differences will therefore affect their co-evolutionary dynamics and the way they respond to changes in their respective environments (e.g. drugs, insecticides).

Acknowledgements

We thank the Institut de Recherche pour le Développement and Centre National de la Recherche Scientifique for financial support. This project was supported by the program PAL+ of the French Ministry of Education and Research. We are grateful to the Kenya Medical Research Institute for their kind help during mosquito collection and people from UR16 (IRD, Montpellier) for their help in characterising mosquito molecular forms.

References

- [1] S. Wright, The genetical structure of populations, *Annals of Eugenics* 15 (1951) 323–354.
- [2] C.D. Criscione, R. Poulin, M.S. Blouin, Molecular ecology of parasites: elucidating ecological and microevolutionary processes, *Molecular Ecology* 14 (2005) 2247–2257.
- [3] P. Awadalla, D. Walliker, H. Babiker, M. Mackinnon, The question of *Plasmodium falciparum* population structure, *Trends in Parasitology* 17 (2001) 351–353.
- [4] D.L. Hartl, S.K. Volkman, K.M. Nielsen, A.E. Barry, K.P. Day, D.F. Wirth, E.A. Winzeler, The paradoxical population genetics of *Plasmodium falciparum*, *Trends in Parasitology* 18 (2002) 266–272.
- [5] S. Gandon, Y. Michalakis, Local adaptation, evolutionary potential and host-parasite coevolution: interactions between migration, mutation, population size and generation time, *Journal of Evolutionary Biology* 15 (2002) 451–462.
- [6] I.M. Hastings, The origins of antimalarial drug resistance, *Trends in Parasitology* 20 (2004) 512–518.
- [7] C. Boete, J.C. Koella, Evolutionary ideas about genetically manipulated mosquitoes and malaria control, *Trends in Parasitology* 19 (2003) 32–38.
- [8] F.G. Razakandrainibe, P. Durand, J.C. Koella, T. De Meeus, F. Rousset, F.J. Ayala, F. Renaud, “Clonal” population structure of the malaria agent *Plasmodium falciparum* in high-infection regions, *Proceedings of the National Academy of Sciences of the United States of America* 102 (2005) 17388–17393.

- [9] J.A. Scott, W.G. Brogdon, F.H. Collins, Identification of single specimens of the *Anopheles-gambiae* complex by the polymerase chain-reaction, *American Journal of Tropical Medicine and Hygiene* 49 (1993) 520–529.
- [10] L. Kamau, T. Lehmann, W.A. Hawley, A.S.S. Orago, F.H. Collins, Microgeographic genetic differentiation of *Anopheles gambiae* mosquitoes from Asembo Bay, Western Kenya: A comparison with Kilifi in coastal Kenya, *American Journal of Tropical Medicine and Hygiene* 58 (1998) 64–69.
- [11] T. Lehmann, N.J. Besansky, W.A. Hawley, T.G. Fahey, L. Kamau, F.H. Collins, Microgeographic structure of *Anopheles gambiae* in western Kenya based on mtDNA and microsatellite loci, *Molecular Ecology* 6 (1997) 243–253.
- [12] T.J.C. Anderson, B. Haubold, J.T. Williams, J.G. Estrada-Franco, L. Richardson, R. Mollinedo, M. Bockarie, J. Mokili, S. Mharakurwa, N. French, J. Whitworth, I.D. Velez, A.H. Brockman, F. Nosten, M.U. Ferreira, K.P. Day, Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*, *Molecular Biology and Evolution* 17 (2000) 1467–1482.
- [13] J. Goudet, Fstat (vers.1.2): a computer program to calculate F-statistics, *Journal of Heredity* 86 (1995) 485–486.
- [14] M. Nei, *Molecular Evolutionary Genetics*, Columbia University Press, New York, 1987.
- [15] B.S. Weir, C.C. Cockerham, Estimating F-statistics for the analysis of population structure, *Evolution* 38 (1984) 1358–1370.
- [16] G. Goudet, T. Raymond, T. De Meeûs, F. Rousset, Testing differentiation in diploid populations, *Genetics* 144 (1996) 1933–1940.
- [17] R.C. Yang, Estimating hierarchical F-statistics, *Evolution* 52 (1998) 950–956.
- [18] J. Goudet, HIERFSTAT, a package for R to compute and test hierarchical F-statistics, *Molecular Ecology Notes* 5 (2005) 184–186.
- [19] F. Rousset, Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance, *Genetics* 145 (1997) 1219–1228.
- [20] W.R. Rice, Analysing tables of statistical tests, *Evolution* 43 (1989) 223–225.
- [21] S.R.G. Nyanjom, H. Chen, T. Gebre-Michael, E. Bekele, J. Shililu, J. Githure, J.C. Beier, G. Yan, Population genetic structure of *Anopheles arabiensis* mosquitoes in Ethiopia and Eritrea, *Journal of Heredity* 94 (2003) 457–463.
- [22] J. Carnahan, L. Zheng, C.E. Taylor, Y.T. Toure, D.E. Norris, G. Dolo, M. Diuk-Wasser, G.C. Lanzaro, Genetic differentiation of *Anopheles gambiae* s.s. populations in Mali, West Africa, using microsatellite loci, *Journal of Heredity* 93 (2002) 249–253.
- [23] G.C. Lanzaro, Y.T. Toure, J. Carnahan, L.B. Zheng, G. Dolo, S. Traore, V. Petrarca, K.D. Vernick, C.E. Taylor, Complexities in the genetic structure of *Anopheles gambiae* populations in west Africa as revealed by microsatellite DNA analysis, *Proceedings of the National Academy of Sciences of the United States of America* 95 (1998) 14260–14265.
- [24] T. Lehmann, W.A. Hawley, H. Grebert, F.H. Collins, The effective population size of *Anopheles gambiae* in Kenya: Implications for population structure, *Molecular Biology and Evolution* 15 (1998) 264–276.
- [25] T. Lehmann, W.A. Hawley, L. Kamau, D. Fontenille, F. Simard, F.H. Collins, Genetic differentiation of *Anopheles gambiae* populations from East and West Africa: Comparison of microsatellite and allozyme loci, *Heredity* 77 (1996) 192–200.
- [26] T. Lehmann, M. Licht, N. Elissa, B.T.A. Maega, J.M. Chimumbwa, F.T. Watsenga, C.S. Wondji, F. Simard, W.A. Hawley, Population structure of *Anopheles gambiae* in Africa, *Journal of Heredity* 94 (2003) 133–147.
- [27] R.S. Waples, A Generalized approach for estimating effective population-size from temporal changes in allele frequency, *Genetics* 121 (1989) 379–391.
- [28] R.S. Waples, Conservation genetics of Pacific salmon. 2. Effective population-size and the rate of loss of genetic variability, *Journal of Heredity* 81 (1990) 267–276.
- [29] T. De Meeus, Y. Lorimier, F. Renaud, Lyme borreliosis agents and the genetics and sex of their vector, *Ixodes ricinus*, *Microbes and Infection* 6 (2004) 299–304.
- [30] Z. Annan, P. Durand, F.J. Ayala, C. Arnathau, P. Awono-Ambene, F. Simard, F.G. Razakandrainibe, J.C. Koella, D. Fontenille, F. Renaud, Population genetic structure of *Plasmodium falciparum* in the two main African vectors, *Anopheles gambiae* and *Anopheles funestus*, *Proceedings of the National Academy of Sciences of the United States of America* 104 (2007) 7987–7992.
- [31] R.E.L. Paul, K.P. Day, Mating patterns of *Plasmodium falciparum*, *Parasitology Today* 14 (1998) 197–202.
- [32] J.L. Cherry, J. Wakeley, A diffusion approximation for selection and drift in a subdivided population, *Genetics* 163 (2003) 421–428.
- [33] W.G. Hill, A. Robertson, The effect of linkage on limits to artificial selection, *Genetical Research* 8 (1966) 269–294.