Inversion polymorphism and incipient speciation in *Anopheles gambiae* s.str. in The Gambia, West Africa

J. H. Bryan[1], M. A. Di Deco[2], V. Petrarca[2] & M. Coluzzi[2]
[1] *Medical Research Council, Fajara, P.O. Box 273, Banjul, The Gambia*
[2] *Istituto di Parassitologia, Università di Roma, Rome, Italy*

Abstract

Analysis of the inversion polymorphism of *Anopheles gambiae* s.str. in The Gambia and surrounding zones of Senegal (Saloum and Casamance) shows, in samples from the central part of the study area, highly significant and temporally stable departures from Hardy-Weinberg equilibrium with deficits of the heterokaryotypes. This situation and the general pattern of karyotype distribution are consistent with the hypothesis of two chromosomally differentiated populations of *A. gambiae* which show partial reproductive isolation and incomplete intergradation in the contact zone.

Introduction

Three malaria vector mosquitoes of the *Anopheles gambiae* complex are present in The Gambia, West Africa: *A. gambiae* s.str., *A. arabiensis* and *A. melas*. These sibling species can be reliably identified by the banding pattern of their polytene chromosomes which shows characteristic fixed inversions in each species as well as intraspecific inversion polymorphisms. Studies on these chromosomal polymorphisms were undertaken in an attempt to understand the relationships between chromosomal variants and environmental heterogeneities. The Gambia is especially suited for such studies as it provides a range of ecological conditions within a relatively small area; brackish and fresh-water environments are in close proximity and, particularly on the North Bank of The Gambia River, riverine areas are found near more arid areas of Sudan savanna.

Preliminary investigations were carried out in 1979 in The Gambia and more detailed studies in 1980 and 1981 when samples were also obtained from some localities in neighbouring areas of Senegal. The aims of the initial study were to determine which inversions were polymorphic, to investigate the geographical distribution of the inversions and the ecological factors influencing this distribution, and to select an area with a high degree of polymorphism for more intensive studies on adult female behaviour and chromosomal polymorphism. These aims were partly modified by the unexpected results obtained in 1979 and 1980 when it was found that in a number of localities in the central area of The Gambia, in samples of *A. gambiae* s.str. obtained from inside houses during the day, the karyotype frequencies were significantly different from those predicted by the Hardy-Weinberg equilibrium. In this paper we present the results of a geographical survey of *A. gambiae* s.str. and the results of longitudinal studies in a locality in which karyotypes were not in agreement with the Hardy-Weinberg expectations. Possible explanations for the Hardy-Weinberg disequilibria are discussed.

Study area

The study area consisted of The Gambia and areas of Senegal which border The Gambia to the

Genetica 59, 167–176 (1982). 0016-6707/82/0593-0167/$2.00.

North and South (Fig. 1). In this region of West Africa the rainfall generally varies from North to South, being higher in the more southern localities. Most of The Gambia lies between the isohyets of 900 mm and 1 250 mm. The collecting sites in North Senegal (Saloum region) receive less than 1 000 mm of rain per year. Considerably wetter conditions prevail in the Casamance region in South Senegal and the collecting sites of Ziguinchor and Belaye lie between the isohyets of 1 500 mm and 1 750 mm. However in the extreme western part of The Gambia the rainfall is higher so that there the isohyets curve North. Rain falls within a single rainy season which usually begins in late May/early June and ends in October/November.

The most important geographical feature of The Gambia is the Gambia River which runs through the centre of the country throughout the country's length of about 400 km. A number of its tributaries also flow through The Gambia. The main river and its tributaries are tidal throughout The Gambia and salt water penetrates at least 160 km from the mouth of the river. The penetration of salt water is influenced by the flow of freshwater in the river, and so is most extensive at the end of the dry season when it may reach as far inland as Kaur and least extensive during the rainy season when it may reach only as far as Farafenni. Large areas of both *Rizophora* and *Avicennia* mangroves occur in areas which are inundated at high tides by brackish water.

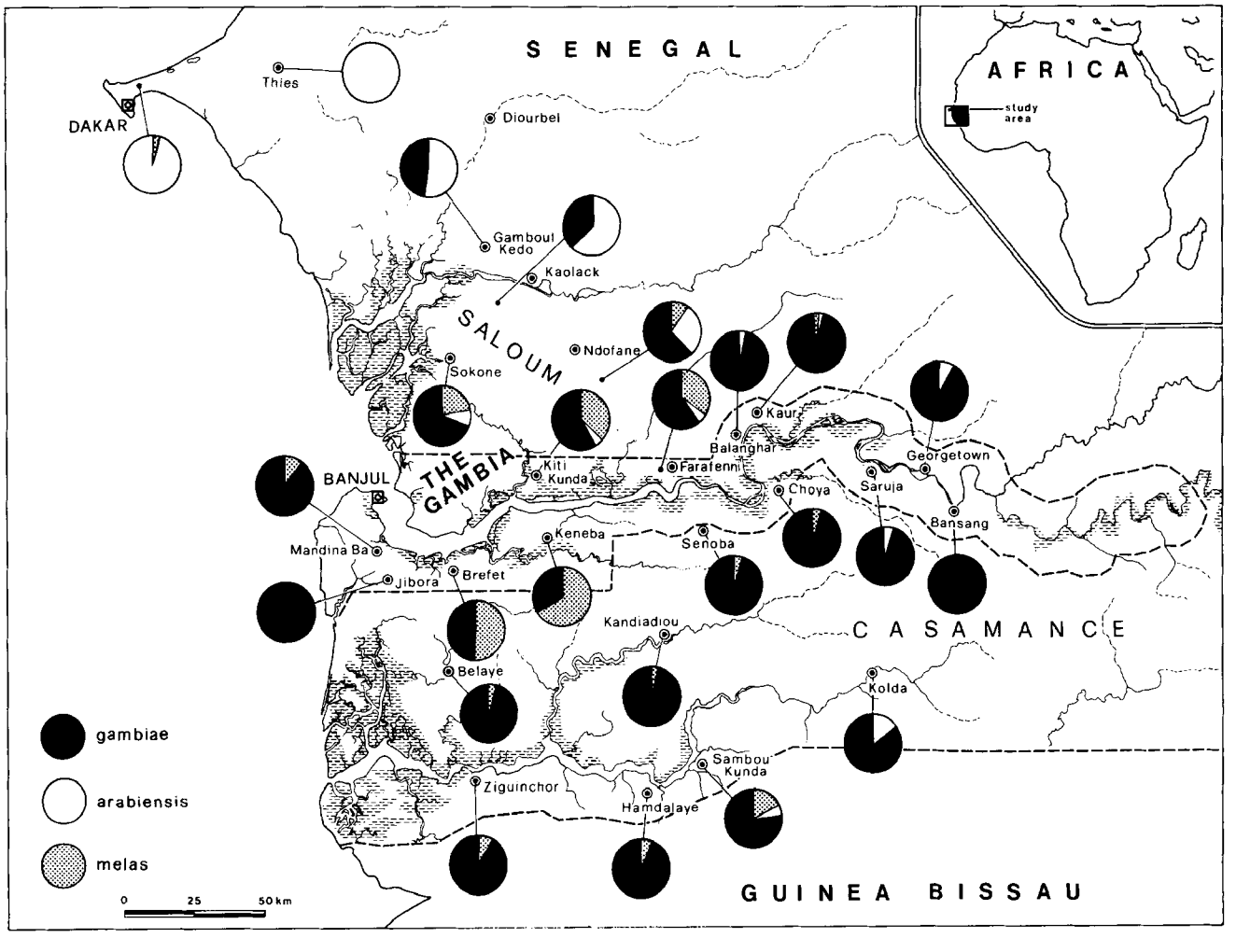

Fig. 1. Distribution of *Anopheles gambiae* s.s., *A. arabiensis* and *A. melas* in indoor resting samples from various localities in The Gambia and bordering areas.

Materials and methods

In 23 localities samples were obtained by daytime indoor resting catches with many specimens collected in mosquito nets. Collections were made in the rainy seasons of 1979, 1980 and 1981. For most localities only one or two samples were obtained and the results from different samples in the same year have been pooled only if no statistically significant difference in the chromosomal frequencies occurred. Numerous samples were collected from Mandina Ba and from the Farafenni area (Ker Madi and Balingho) where longitudinal studies were carried out during the 1980 and 1981 rainy seasons for a study on chromosomal polymorphism in *A. melas* and *A. gambiae* s.str. and for an investigation into the behaviour of the mosquitoes carrying different karyotypes. In these two areas collections were also made from animal shelters, from night biting catches using human bait inside and outside houses, and from calves outside. Mosquitoes collected during the night biting catches were blood fed to repletion within one hour of capture.

After capture, the mosquitoes were stored in the collecting cups in an insulated container until the late afternoon. By this time, at the ambient temperature, the ovaries had reached the most suitable stage for chromosomal scoring. They were fixed in one part glacial acetic acid and two or three parts ethanol immediately after being killed, held for 24 hours at the ambient temperature and then stored at 4 °C or -20 °C. The chromosomes, after preparation following the technique of Hunt (1973) were scored according to the nomenclature of Coluzzi *et al.* (1979). The percentage of unreadable preparations was less than 5%.

The karyotype frequencies in each sample were tested for deviation from the Hardy-Weinberg Law using Wright's F statistic (F). Following Brown (1970) [$F = 4a_1a_3 - a_2^2/(2a_1 + a_2)(2a_3 + a_2)$] where a_1 and a_3 are the frequencies of the two homozygous classes and a_2 the frequency of the heterozygote. There are significant ($P < 0.05$) departures from the Hardy-Weinberg expectations if $|F| > 1.96/\sqrt{N}$. $F > 0$ indicates a deficiency of heterozygotes and $F < 0$ an excess of heterozygotes. Its bias is negligible for samples above 20.

Chromosomal polymorphism was also studied in three colonies of *A. gambiae* s.str. from The Gambia: G3, initially colonized in the London School of Hygiene and Tropical Medicine from material from Georgetown, BAN, a colony from females collected at Bansang and ZIG, a colony from specimens collected at Ziguinchor.

Results

The distribution of *A. gambiae* s.str., *A. arabiensis* and *A. melas* in the indoor resting samples obtained from the study area is illustrated in Figure 1. *A. gambiae* s.str. occurs in all localities except those at the northern limit of the study area where *A. arabiensis* is dominant. The latter species is also present in the Saloum region, in the North Bank and inland areas of The Gambia, and in inland localities of the Casamance such as Kolda. *A. melas* is common in areas where the river systems are tidal and brackish and are bordered by mangroves. This species penetrates as far inland as Choya and Kaur.

In a number of localities only a single sample was obtained and as the species densities, particularly in the case of *A. melas,* are subject to large, not parallel fluctuations (Bryan, 1982), the relative frequencies diagrammatically reported in Figure 1 should not be taken as representative of the mean proportion of the three species in the area. Moreover, as *A. melas* is more exophilic and zoophilic than both *A. gambiae* and *A. arabiensis* it is also likely to be under-represented in collections from houses. For example at Mandina Ba only 3% of indoor resting samples, but 30% or more of samples from cowsheds were *A. melas.*

Within the study area *A. gambiae* is polymorphic for three common independent inversions on the second chromosome: 2Rb, 2Rd and 2La (Fig. 2). Among other inversions, which are observed sporadically in the area, the 2Rbc is the least rare reaching in some localities frequencies between 1

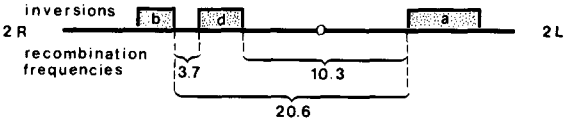

Fig. 2. Diagrammatic representation of the relative lengths and positions of chromosome 2 inversions in *Anopheles gambiae* s.str. The recombination frequencies between the inversions are also shown (data from Di Deco *et al.*, 1980a).

and 2%. This arrangement has been scored as 2Rb on which it is based (Coluzzi *et al.*, 1979).

The variations of the frequencies of the three common inversions between localities are shown in Figure 3 and in Table 1. The frequencies of 2Rb and 2La are very similar in each locality and vary in a parallel way. In the western part of the study area they show a clinal variation from South to North, occurring with a frequency of only 20% or less in Ziguinchor, increasing to around 30% in the Brefet-Jibora area, to 50% at Sokone and to above 80% at Gamboul Kedo. The frequencies of 2Rd vary in an opposite way. Thus it is uncommon in most northern areas and prevalent in southern ones. However the frequencies of all the inversions also show changes in a west-to-east direction. In The Gambia 2Rd is prevalent in the western localities, is common as far east as the Farafenni-Kaur area but becomes very rare in the Georgetown, Saruja and Bansang samples. Again where 2Rd decreases 2Rb and 2La increase. Similar changes in a west-to-east direction occur also in the Casamance.

The changes in the frequencies of 2Rb and 2La in the western part of the study area can be interpreted as a cline related to increased aridity. It has already been shown in Nigeria (Coluzzi *et al.*, 1979) that the carriers of 2Rb and 2La are more prevalent in arid areas than the carriers of the standard chromosome whereas 2Rd (alone and not associated with 2Rbc) appears to be widespread in humid savanna or for-

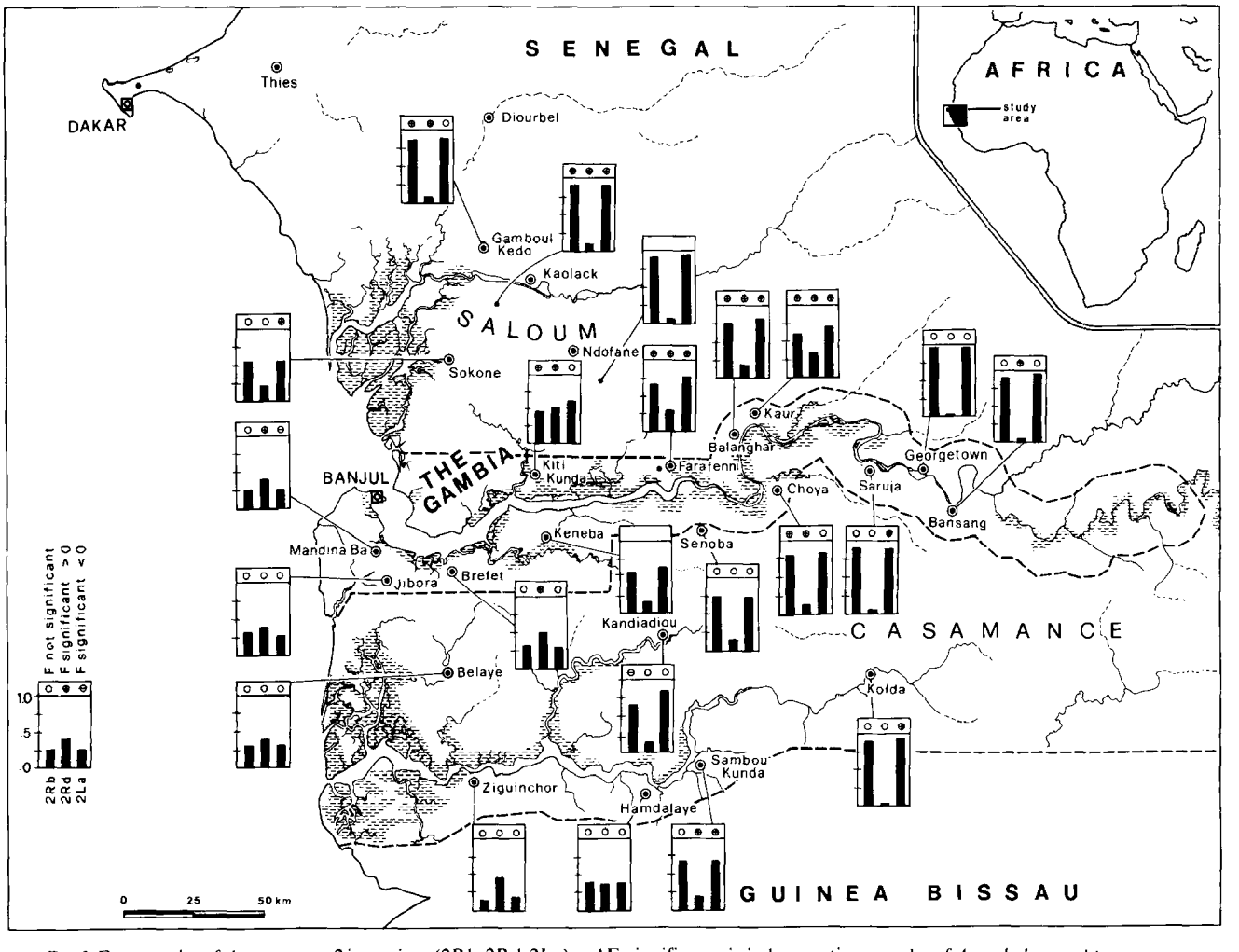

Fig. 3. Frequencies of chromosome 2 inversions (2Rb, 2Rd, 2La) and F, significance in indoor resting samples of *Anopheles gambiae* s.str. from various localities in The Gambia and bordering areas (see Table 1 and the text for further details).

est-savanna localities (Coluzzi *et al.*, 1979 and unpublished data).

Much more puzzling are the changes in the inversion frequencies in a west-east direction. The rainfall throughout this area is similar, although temperature conditions become more extreme in inland regions. However, the environmental climatic changes are gradual and therefore do not satisfactorily explain the abrupt changes which occur east of Kaur in the frequencies of all the inversions and in particular, the rapid decrease in 2Rd.

That such chromosomal variations cannot be easily interpreted as clinal frequency changes in a single panmictic population is demonstrated when the karyotype frequencies are tested against Hardy-Weinberg expectations (see F-values in Table 1). While in samples from the extreme limits of the 'cline' such as Jibora, Ziguinchor, Georgetown and Kolda the karyotype frequencies are not significantly different from H-W expectations, most of the other samples are significantly different for one or more inversion systems, particularly in the 2Rb and 2Rd systems. These differences are greatest in areas in The Gambia in which a fairly high frequency of 2Rd coexists with high frequencies of 2Rb and 2La. In all such cases F is positive and generally significant indicating a large deficit of heterozygotes.

An examination of the association between 2Rb and 2Rd karyotypes in the Balanghar and Farafenni samples (Tables 2 and 3) sheds some light on the problem presented by this phenomenon. In the Balanghar sample only two of the four classes of the expected 2R homokaryotypes are represented, namely $+^b/+^b$–d/d and b/b–$+^d/+^d$, while $+^b/+^b$–$+^d/+^d$ and b/b–d/d are completely absent

Table 1. Frequencies of chromosome-2 inversions in indoor resting samples of *Anopheles gambiae* s.str. from various localities in The Gambia and bordering areas in Senegal. F is a measure of Hardy-Weinberg deviation; value is significant (*) if $|F| > 1.96/\sqrt{N}$ (see the text for further details).

Locality	Date	N	$\frac{1.96}{\sqrt{N}}$	2Rb	F	2Rd	F	2La	F
Senegal-Saloum									
Ndofane	Aug '80	9	0.65	0.8889	0.12	0.0556	-0.06	0.9444	-0.06
Kaolack	Aug–Oct '80	41	0.31	0.8780	0.54*	0.0976	0.45*	0.8902	0.38*
Gamboul Kedo	Aug–Oct '80	60	0.25	0.8583	0.52*	0.0833	0.56*	0.8833	0.03
Sokone	Aug–Oct '80	35	0.33	0.5714	0.17	0.2000	-0.25	0.5571	0.36*
The Gambia									
Bansang	Sept–Oct '79	50	0.28	0.9100	0.15	0.700	0.54*	0.9479	-0.05
Georgetown	Aug–Nov '79	37	0.39	0.9324	-0.07	0.0135	-0.00	0.9590	-0.04
Saruja	Oct '80	105	0.19	0.8714	0.05	0.0429	0.19	0.9190	0.30*
Choya	Nov '80	197	0.14	0.7716	0.43*	0.1726	0.61*	0.8528	0.07
Kaur	Oct '80	170	0.15	0.6176	0.51*	0.3294	0.71*	0.7118	0.23*
Balanghar	Oct '80	240	0.13	0.7333	0.42*	0.2062	0.63*	0.8229	0.16*
Farafenni	Sept '80	88	0.21	0.6364	0.51*	0.3068	0.68*	0.7500	0.33*
Farafenni	Oct '81	530	0.09	0.7321	0.42*	0.2085	0.57*	0.8226	0.31*
Keneba	Oct '80	16	0.49	0.5625	0.75*	0.1562	0.76*	0.5938	0.09
Kiti Kunda	Oct '80	71	0.23	0.4296	0.51*	0.4859	0.63*	0.5845	0.22
Brefet	Sept–Oct '79	215	0.13	0.2907	0.11	0.4953	0.17*	0.3209	-0.07
Jibora	Sept–Oct '79	446	0.09	0.3038	0.02	0.3722	0.07	0.2332	-0.07
Jibora	Sept '80	94	0.20	0.3404	0.05	0.3245	-0.14	0.2660	0.07
Mandina Ba	Aug–Dec '80	1 192	0.06	0.2655	-0.03	0.4128	0.07*	0.2693	-0.13*
Senegal-Casamance									
Kolda	Jul '81	96	0.20	0.8958	-0.12	0.0208	0.02	0.9583	0.22*
Sambou Kunda	Jul '81	58	0.26	0.6897	0.20	0.2069	0.37*	0.6983	0.47*
Senoba	Oct '81	42	0.30	0.7738	0.12	0.1548	-0.00	0.7738	0.25
Kandiadiou	Jul '81	43	0.30	0.6512	-0.54*	0.1396	0.03	0.8256	0.11
Hamdalaye	Jul '81	76	0.22	0.4013	0.04	0.3882	0.08	0.3947	0.12
Ziguinchor	Sept '80	102	0.10	0.1618	0.02	0.4706	0.02	0.2059	0.04
Ziguinchor	Aug '81	117	0.18	0.1496	0.04	0.4402	0.13	0.2051	0.15
Belaye	Sept '80	62	0.25	0.2984	0.11	0.4193	0.21	0.2984	0.19

although the expected numbers are 25.62 and 22.19 respectively. This situation could be interpreted in terms of a strong or complete linkage disequilibrium between the 2Rb and 2Rd systems in agreement with the relatively low (less than 4%) recombination frequency between the two inversion systems (Di Deco *et al.*, 1980a). However, in this case one would also expect that most, if not all, the heterozygotes would be double 2R heterokaryotypes ($+^b/b$–$+^d/d$) but the other four classes of heterokaryotypes are all present with $+^b/b$–$+^d/+^d$ occurring even more frequently than the double heterokaryotype. The total number of heterozygotes is only 62 compared to 31 and 142 in the two homozygous classes. Once again, there is an obvious deficit of the expected heterozygotes, similar to that seen when the inversions are considered individually. Moreover only 27% (17/62) of the heterozygotes are double heterozygotes. Thus the figures for the association of the inversion systems cannot be interpreted as a linkage disequilibrium in a panmictic population since these figures would indicate dramatic selection pressures acting not only against the 2R recombinant homokaryotypes $+^b/+^b$–$+^d/+^d$ and b/b–d/d but also against the double heterokaryotype $+^b/b$–$+^d/d$.

Table 2. Numbers of observed (o) and expected (e) associations between 2Rb and 2Rd inversion karyotypes in the sample of *Anopheles gambiae* s.str. from Balanghar. Expected figures were calculated from the marginal totals.

2Rd karyotypes		2Rb karyotypes $+^b/+^b$	2Rb karyotypes $+^b/b$	2Rb karyotypes b/b	Totals
$+^d/+^d$	o	0	30	142	172
	e	25.62	37.33	109.06	
	$\frac{(o-e)^2}{e}$	25.62	1.44	9.95	
$+^d/d$	o	4	17	7	28
	e	4.17	6.08	17.75	
	$\frac{(o-e)^2}{e}$	0.007	19.61	6.51	
d/d	o	31	4	0	35
	e	5.21	7.60	22.19	
	$\frac{(o-e)^2}{e}$	127.66	1.71	22.19	
Totals		35	51	149	235

$\chi^2_4 = \Sigma \frac{(o-e)^2}{e} = 214.70\ (P < 0.001)$

Table 3. Number of observed (o) and expected (e) associations between 2Rb and 2Rd inversion karyotypes in the sample of *Anopheles gambiae* s.str. from Farafenni (Ker Madi). Expected figures were calculated from marginal totals. Four specimens carriers of the 2Rbc arrangement at the heterozygous state have not be considered.

2Rd karyotypes		2Rb karyotypes $+^b/+^b$	2Rb karyotypes $+^b/b$	2Rb karyotypes b/b	Totals
$+^b/+^b$	o	4	61	315	380
	e	59.24	86.69	234.07	
	$\frac{(o-e)^2}{e}$	51.51	7.61	27.98	
$+^d/d$	o	9	55	9	73
	e	11.38	16.65	44.97	
	$\frac{(o-e)^2}{e}$	0.50	88.33	28.77	
d/d	o	69	4	0	73
	e	11.38	16.65	44.97	
	$\frac{(o-e)^2}{e}$	291.75	9.61	44.97	
Totals		82	120	324	526

$\chi^2_4 = \Sigma \frac{(o-e)^2}{e} = 551.03\ (P < 0.001)$

Karyotype distribution is similarly unbalanced in the Farafenni sample (Table 3). There are only four carriers of the standard 2R homokaryotype and none of the doubly inverted homokaryotype although the expected numbers are 59.24 and 44.97. There are 69 carriers of $+^b/+^b$–d/d and 315 carriers of b/b–$+^d/+^d$ while the total number of heterozygotes is only 138 of which only 40% are double heterozygotes. This same pattern of karyotype distribution occurs in most localities from the central part of the study area such as Keneba, Kiti Kunda, Kaur and Choya. All the samples from these localities demonstrate very high levels of Hardy-Weinberg disequilibrium. Moreover, one or both of the 2R homokaryotypes exceed the expectation in practically all samples except some from the limits of the study area.

The unbalance of karyotype distribution as shown in Table 2 and 3 is not dependent on the

collection method. Samples collected at Farafenni with night biting catches on man outdoors and indoors, and on calves outdoors, show some changes in the frequency of certain chromosomal arrangements (which will be considered in detail in a separate paper) but the general pattern of karyotype distribution is the same recorded in the samples of *A. gambiae* collected resting indoors in the same locality. Accordingly the excess of $+^{b}$–d and b–$+^{d}$ homokaryotypes and the deficit of the corresponding double heterokaryotype appears to be a real phenomenon, at least for the female biting population.

Further evidence for genetic heterogeneity in most of the samples, is provided by the association of the two alternative 2R homokaryotypes $+^{b}/+^{b}$–d/d and b/b–$+^{d}/+^{d}$ with 2La karyotypes. These should occur in similar frequencies in the two homokaryotype subsamples of each sample particularly considering that recombination frequencies are estimated to be 10% between 2Rd and 2La and 26% between 2Rb and 2La (Di Deco *et al.*, 1980a). However, the frequency of 2La is generally much higher among b–$+^{d}$ 2R homokaryotypes than among $+^{b}$–d 2R homokaryotypes and this is particularly significant in the samples from the central part of the study area as shown in Table 4. Moreover, while the frequencies of 2La karyotypes do not depart significantly from Hardy-Weinberg expectations in each of the 2R homokaryotype subsamples (N_1 and N_2) with most F values negative indicating an excess of heterozygotes, the 2La frequencies calculated from the sum of the two subsamples ($N_1 + N_2$) are nearly coincident with those of the total sample (N_t) and both show a significant deficit of heterozygotes. This is fully consistent with the hypothesis that the deficit of heterozygotes in the total sample is a consequence of the mixing of two populations with different inversion frequencies (Wahlund's principle). Thus, assuming the 2R homokaryotypes b–$+^{d}$ and $+^{b}$–d as reliable markers for sampling population 1 and population 2 respectively, the relative proportions of these two populations can be estimated by using 2La frequencies. If p_t is the frequency of 2La in the total sample and p_1 and p_2 are the 2La frequencies in population 1 and 2 respectively, the frequency of population 2 would be $p_t - p_1 / p_2 - p_1$ (Wallace, 1968) as shown in the last column of Table 4.

Additional support for the hypothesis that there are two populations of *A. gambiae* which partially intergrade in the central part of the study area is provided by the data from the longitudinal study carried out at Farafenni. The unbalance of karyotype distribution does not appear as a transitory phenomenon: it is recorded in this locality in 1980 and 1981 and it seems fairly stable during the 1981 rainy season in spite of changes in the relative proportion of the two populations, with population 2 increasing particularly during September, October and November (Table 5).

The data obtained from laboratory colonies are not indicative of the presence of any mechanism of reproductive isolation between the chromosomal types characterizing the two populations. Colonies founded from female carriers of $+^{b}$–d 2R homokaryotypes (ZIG) and of b–$+^{d}$ 2R homokaryotypes (BAN) are crossed and backcrossed without difficulty. The double heterokaryotypes $+^{b}/$b–$+^{d}$/d, which occur less frequently than expected in nature,

Table 4. Frequency (p) of the 2La arrangement in *Anopheles gambiae* s.str. carriers of *b*–$+^{d}$ and $+^{b}$–*d* 2R homokaryotypes (p_1 and p_2) in different samples from the central part of the study area. The data from the total samples (p_t), already reported in Table 1, are included to allow easier comparisons with the data obtained from the sum of the two homokaryotypes subsamples ($p_1 + p_2$). Significant F values are marked by an asterisk. (See the text for further details).

Sample	*b*–$+^{d}$ homokaryotypes			$+^{b}$–*d* homokaryotypes			Both homokaryotypes			Total sample			$p_t - p_1$
	N_1	p_1	$\hat{F}$	N_2	p_2	$\hat{F}$	$N_1 + N_2$	p_{1+2}	$\hat{F}$	N_t	p_t	$\hat{F}$	$p_2 - p_1$
Kiti Kunda	22	0.8864	–0.13	26	0.4423	–0.17	48	0.6458	0.09	71	0.5845	0.22	0.6798
Farafenni 1980	45	0.9444	–0.06	21	0.4286	0.03	66	0.7803	0.34*	88	0.7500	0.33*	0.3779
Farafenni 1981	306	0.9346	–0.07	63	0.3571	0.14	369	0.8360	0.36*	530	0.8226	0.31*	0.1939
Balanghar	142	0.9120	–0.01	31	0.4677	–0.10	173	0.8324	0.17*	235	0.8255	0.14*	0.1947
Kaur	83	0.8855	–0.01	41	0.3658	–0.26	124	0.7137	0.19*	169	0.7130	0.23*	0.3319
Choya	125	0.9000	–0.11	14	0.3214	0.51	139	0.8417	0.24*	196	0.8546	0.08	0.1692

Table 5. Frequency (p) of the 2La arrangement in *Anopheles gambiae* s.str. carriers of $b^-/+^d$ and $+^b/-^d$ 2R homokaryotypes (p_1 and p_2) in samples collected at Farafenni (Ker Madi) during different periods of the 1981 rainy season. (See Table 4 and the text for further details).

Sample	$b^-/+^d$ homokaryotypes			$+^b/-^d$ homokaryotypes			Both homokaryotypes			Total sample			$p_t - p_1$
	N_1	p_1	$\hat{F}$	N_2	p_2	$\hat{F}$	N_1+N_2	p_{1+2}	$\hat{F}$	N_t	p_t	$\hat{F}$	$p_2 - p_1$
July	16	0.9062	-0.10	-	-	-	-	-	-	22	0.8409	-0.19	0.0721
September	71	0.9155	-0.09	16	0.4688	-0.13	87	0.8333	0.13	120	0.8292	0.15	0.1930
October I	235	0.9362	-0.07	43	0.3837	0.07	278	0.8507	0.31*	383	0.8356	0.28*	0.1514
October II	71	0.9296	-0.08	20	0.3000	0.29	91	0.7912	0.47*	147	0.7891	0.39*	0.2232
November I	33	0.9242	-0.08	25	0.3000	0.14	58	0.6552	0.47*	83	0.6687	0.37*	0.4093
November II	29	0.9310	-0.07	37	0.3649	0.13	66	0.6136	0.39*	84	0.6131	0.37*	0.5616
December	9	0.9444	-0.06	13	0.3461	-0.19	22	0.5909	0.25	27	0.6667	0.33	0.4641

are found to be not only viable and fertile in the laboratory but even heterotic since their frequency generally exceeds the expected levels in cage populations. Significant excess of heterokaryotypes is commonly observed in the G3 colony polymorphic for inversions 2Rb, 2Rbc, 2Rd and 2La (Di Deco *et al.*, 1980b). Furthermore, while the linkage disequilibrium may be maintained between the 2R inversions systems b and d (recombination frequency lower than 4%), random association prevails between 2Rd and 2La which have a recombination frequency of 10% (Di Deco *et al.*, 1980a).

Discussion

Analysis of the chromosomal polymorphism of *A. gambiae* s.str. in the study area indicates that there are two main chromosomal types which coexist in some localities in variable proportions. One chromosomal type is characterized by a high frequency of 2Rb and 2La arrangements, while 2Rd is very rare or absent. This chromosomal type is most frequent in areas of The Gambia east of Kaur, in the eastern part of the Casamance such as Kolda and in North Senegal. Its frequency decreases rapidly in the western part of the study area. The second chromosomal type has a lower frequency of arrangements 2Rb and 2La and a high frequency of 2Rd arrangement. It is common, or the only one present in the western Casamance samples (Belaye, Ziguinchor and Hamdalaye) and is the most common one in the western part of The Gambia (Mandina Ba, Jibora and Brefet) but becomes rarer in the East and is practically absent east of Kaur. It decreases in frequency in North Senegal although it is still present in Kaolack and Gamboul Kedo.

Assuming that we are dealing with a single panmictic population, it appears at first that there is a strong linkage disequilibrium between the three inversion systems of chromosome 2. The favoured linkage would be for the two alternative arrangements $\underline{+^b\text{-}d\text{-}+^a}$ and $\underline{b^-\text{-}+^d\text{-}a}$ since the corresponding homokaryotypes are present more frequently than expected. However, in areas where both chromosomal types are common there is evidence that heterokaryotypes between them occur less frequently than expected, or are absent. Moreover, each of the inversion systems shows a Hardy-Weinberg disequilibrium generally with a highly significant deficit of heterokaryotypes.

Considering the laboratory evidence that there are no genetic barriers between the two prevailing homokaryotypes and assuming random mating between them, we should conclude that the deficit or absence of expected heterokaryotypes and recombinant homokaryotypes in the adult female samples is because mosquitoes bearing such karyotypes have such low fitness in nature that they die in the pre-imaginal stages or soon after emergence. However the assumption of such extreme differences in karyotype fitness and in particular the assumption of a lower fitness of the heterokaryotypes would imply a very unstable equilibrium, which is problematic to hypothesize unless it is a transitory phenomenon leading to isolation or to disruption of the prevailing inversion associations. In contrast, the longitudinal study at Farafenni indicates that this phenomenon is not only widespread but temporally stable.

A more acceptable hypothesis to explain the pattern of karyotype distribution recorded in the central part of the study area is that we are dealing with a zone of incomplete intergradation between two populations of *A. gambiae* s.str. Their typical chromosomal constitutions closely correspond to the two chromosomal types discussed above; their most characteristic homokaryotypes are b/b-$+^d/+^d$-a/a for population 1 and $+^b/+^b$-d/d-$+^a/+^a$ for population 2. Inversions 2Rb and 2La are polymorphic in both populations but have higher frequencies in population 1. Inversion 2Rd appears to be typical of population 2, where it reaches frequencies around 50%, while it is very rare or absent in population 1. Where the two populations are sampled in the same locality the mixed sample shows apparent linkage disequilibria between the three inversions and Hardy-Weinberg disequilibria for each of them with deficits of heterozygotes as expected from Wahlund's principle.

While the two-populations hypothesis provides a coherent explanation of the pattern of chromosomal variations observed in the study area, it also raises many questions, the most important of which concerns the mechanism of isolation between population 1 and 2 and their taxonomic status.

The two populations are sympatric as adult females at their biting and resting sites but they could be separated at their larval and mating sites. However, this isolation is unlikely to be maintained by purely extrinsic factors. Even assuming that the mixed samples consist of migrants from ecologically different areas where they are subjected to highly diversifying selection pressures, the genetic heterogeneity would be very unstable in successive generations unless intrinsic mechanisms of reproductive isolation are operating. Particularly during the rainy season, 'typical' *gambiae* larval habitats are abundant and often produced within the villages by man's activities: in such conditions, if the choice of oviposition sites is uniform among females of the two populations, these should rapidly become mixed, or, at least, the degree of chromosomal divergence should be progressively reduced in the intergradation zone. The data obtained from the longitudinal study in Farafenni show that the two populations maintain their genetic identity towards the end of the rainy season; accordingly, intrinsic mechanisms of reproductive isolation are likely to be involved and should act at the pre-mating level since the laboratory evidence excludes post-mating barriers. Pre-mating reproductive isolation could result from differences in oviposition-site selection and/or in mating behaviour.

The geographical patterns of distribution of the two populations are clearly indicative of an ecotypic differentiation. Population 1 is the prevalent population or the only one present in fresh-water inland zones of the study area. It can be regarded as a typical dry-savanna population of *A. gambiae* s.str. since its characteristic chromosome arrangement 2Rb-2La is widespread in most Sudan and Sahel Savanna zones both in West and East Africa (Coluzzi *et al.*, unpublished data). Population 2 is prevalent in the more humid, coastal and southern zones of the study area and its distribution is closely related to the range of *A. melas* and to the presence of brackish waters. Both in The Gambia and in Casamance carriers of 2Rd are able to expand successfully inland but only within areas which are inundated by salt water at least seasonally. Moreover population 2 seems completely intergraded with typical forest populations, monomorphic for the 2R standard arrangement, which are present south of the Casamance along the coast (Coluzzi *et al.*, unpublished data from Guinea Bissau and Liberia).

The biogeographical picture, as stated above, appears to be consistent with the assumption of larval competition between the two populations and of higher ability of population 2 to exploit brackish waters for oviposition and/or larval development. Population 2 would not spread in fresh water inland zones because of competition from population 1, which is presumably more adapted to the prevailing savanna environmental conditions. However, this competition would be overcome, in zones under tidal influence, by the ability of population 2 to colonize breeding places with water of a low salt content, avoided by both *A. melas* and population 1. The seasonal changes in the proportion of the two populations, observed at Farafenni during the 1981 rainy season, fully support this interpretation. The brackish environments are absent or rare in the dry season or at the beginning of the rainy season but become very common as the rains progress and saline areas are flooded with fresh water. Human activity in connection with rice cultivation provides still more of such brackish breeding places. Accordingly, an increase of the

relative frequency of population 2 is expected during the second part of the rainy season just as documented by the longitudinal study at Farafenni.

The available data do not allow definite conclusions to be drawn as to the type(s) of premating mechanism of isolation acting between the two populations. Assuming a close association of mating sites with larval habitats, the differential ability to exploit brackish waters for oviposition and/or larval development could result at least in partial reproductive isolation. However the possible existence of differences in mating behaviour should also be taken into account.

With respect to taxonomic status, two populations which coexist without losing their identities should be regarded as at least incipient species (Mayr, 1970). The extent of hybridization occurring in the contact zones, although difficult to evaluate precisely, appears sufficient to allow progressive introgression and complete intergradation between the two gene pools. This should result in clinal frequency changes between the alternative chromosomal types and no significant deviations from Hardy-Weinberg equilibrium. The lack of such complete intergradation suggests that the hybrid genotypes are unfavoured in nature in spite of their heterosis in laboratory colonies.

Better taxonomic definitions of the two populations will depend on further data on their ecological and genetical differentiation (including data on genetic distances) and on their geographical distribution in the more humid zones south of the study area. It will be particularly important to ascertain whether the two taxa are both in contact, through a chain of intergrading populations, with the forest *A. gambiae* monomorphic for the standard arrangement.

Ackowledgements

We gratefully record our thank to Messieurs Pierre Sambou, Bakary Camara and Afansu Colley and Mrs Ida Secka for their help in collecting the mosquitoes. We would also like to thank Mr T. Wilkes for the Saruja sample, Dr J. Vercruysse for samples from the Dakar area and Mr M. Sarr for the Thies area samples. The processing of the specimens for chromosomal examinations was carried out with the excellent technical assistance of Mr G. Petrangeli who also helped in reading the slides. Misses A. Sforzini, M. Saliola, B. Sampaolese and M. Faiola helped in preparing the specimens.

This work was supported by the Italian National Research Council (grant No 79.01973.04), the British Medical Research Council and the World Health Organization.

References

Bryan, J. H., 1979. Observations on the member species of the Anopheles gambiae complex in The Gambia, West Africa. Trans. r. Soc. trop. Med. Hyg. 73: 463–466.

Bryan, J. H., 1982. Anopheles gambiae and An. melas at Brefet, The Gambia, and their role in malaria transmission. Ann. trop. Med. Parasit. (in press).

Brown, A. H. D., 1970. The estimation of Wright's fixation index from genotypic frequencies. Genetica 41: 399–406.

Coluzzi, M., Sabatini, A., Petrarca, V. & Di Deco, M. A., 1979. Chromosomal differentiation and adaptation to human environments in the Anopheles gambiae complex. Trans. r. Soc. trop. Med. Hyg. 73: 483–497.

Di Deco, M. A., Petrarca, V., Villani, F. & Coluzzi, M., 1980a. Recombination and linkage disequilibria between chromosome 2 inversion in Anopheles gambiae s.s. Proc. Third European Multicolloquium of Parasitology, Cambridge. Abs. 87.

Di Deco, M. A., Petrarca, V., Villani, F. & Coluzzi, M., 1980b. Polimorfismo cromosomico da inversioni paracentriche ed eccesso degli eterocariotipi in ceppi di Anopheles allevati in laboratorio. Parassitologia 22: 304–306.

Hunt, R. H., 1973. A cytological technique for the study of Anopheles gambiae complex. Parassitologia 15: 137–139.

Mayr, E., 1970. Populations, species and evolution. Belknap Press, Harvard.

Wallace, B., 1968. Topics in population genetics. Norton, New York.

Received 21.10.1981 Accepted 18.6.1982.