Swarming and mating systems in natural populations of *Anopheles gambiae* M and S forms: review of 4 years survey in rural areas of sympatry, Burkina Faso (West Africa)

Sawadogo P.S.[1], Namountougou M.[1], Toé KH.[1], Rouamba J.[1], Maïga H.[1], Ouédraogo K.R.[1], Baldet T.[2], Gouagna L.C.[3], Kengne P.[4], Simard F.[4], Costantini C.[4], Gibson G.[5], Diabaté A.[1], Lees R.S.[6], Gilles J.[6], Dabiré K.R.[1*]

[1]Institut de Recherche en Sciences de la Santé (IRSS)/Centre Muraz, BP 390 Bobo-Dioulasso, Burkina Faso

[3]Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), UMR CMAEE, Montpellier, France.

[3]Institut de Recherche pour le Développement (IRD), UMR 224- MIVEGEC, Montpellier, France IRD/UR016-CRVOI, 2 rue Maxime Rivière 97490, Sainte Clotilde, Ile de la Réunion

[4]Institut de Recherche pour le Développement (IRD), UMR 224- MIVEGEC, Montpellier, France

[5]Natural Resources Institute /University of Greenwich, Central Avenue Chatham Maritime Kent, ME44TB, UK

[6]Insect Pest Control Laboratory, FAO/IAEA, A2444, Seibersdorf, Vienna, Austria

*corresponding author: dabire_roch@hotmail.com
                           Tel: +226 70 73 90 69
                           Fax: +226 20 97 04 57

Abstract

The swarming and mating systems of natural populations of *An. gambiae* M and S forms were investigated through longitudinal surveys conducted between July 2006 and October 2009 in Soumousso and Vallée du Kou (VK7), two rural areas of south-western Burkina Faso where these forms are sympatric. In both sites, the majority of swarms were recorded above visual markers localized within human habitats. In Soumousso, a wooded area of savannah, 108 pairs caught *in copula* from 205 swarms were sampled; in VK7, a rice growing area, 491 couples from 250 swarms were sampled. In neither site was any spatial segregation observed between the swarm sites used by the two forms of *An. gambiae,* which shared many of their visual markers. Furthermore, mixed swarms were collected annually in frequencies varying from one

site to another, though no mixed inseminations were recorded, corroborating the low hybrid rate previously reported in the field. The occurrence of inter-specific mate-recognition mechanisms, which allow individuals to avoid hybridization, is discussed.

Key-words: *Anopheles gambiae* s.s, Mating, Swarms, Burkina Faso

1. Introduction

Anopheles gambiae s.s. Giles, formerly considered as a single species, was first reported to be accumulating genetic heterogeneity 30 years ago, suggesting the development of genetic reproductive isolation. This mosquito species is highly polymorphic and has been subdivided into five chromosomal forms that differ in their inversion arrangements (Coluzzi et al. 1985; Touré et al. 1998; Coluzzi et al. 2002). These chromosomal forms appear to be genetically isolated in the field, presumably through prezygotic barriers as viable and fertile hybrids have been obtained in the laboratory (Persiani et al. 1986). Cytogenetic analysis, however, is not the most accurate way to evaluate the degree of hybridisation between molecular forms because of the presence of cryptic 'heterokaryotypes', which are impossible to identify with confidence, and also because of the adaptive nature of inversions exposed to strong selection pressure (della Torre et al. 2001; Coluzzi et al. 2002; Wondji et al. 2002). Previous studies, based on molecular markers such as X-linked ribosomal DNA, grouped these chromosomal forms into two entities referred to as M and S molecular forms (Favia et al. 2001). Although these molecular forms are able to hybridise in the laboratory, M/S hybrids are very rare in nature (della Torre et al. 2001, 2005). The issue of reproductive isolation between these molecular forms has created much debate (Coluzzi et al. 2002; della Torre et al. 2002; Caputo et al. 2008 Lawniczak et al. 2010) as the assessment of reproductive isolation varies considerably across their geographical distribution (della Torre et al. 2001; Tripet et al. 2001; Oliveira et al. 2008; Diabaté et al. 2009; Caputo et al. 2011), and we do not yet know how reproductive barriers between these molecular forms function. In fact a cryptic group of *An. gambiae* was only recently found in Burkina Faso (Riehle et al. 2011). Mating systems based on aerial male aggregations that function as encounter sites for mate-searching females have evolved repeatedly in various groups of insects (Sullivan 1981; Cooter 1989). Swarms of most species are composed of males from which

3

females acquire a mate before leaving the swarm *in copula*. Charlwood and Jones (1979) suggested that male *An. gambiae* avoid contact with inter-specific partners mainly by swarming at different heights above a marker, and secondarily by swarming at different times. If this were true, one would expect different swarming behaviour between M and S forms, thereby reducing the chances of contact between males and females of the two forms in a sympatric area. In support of this hypothesis, Tripet et al. (2001) observed assortative mating between the forms in the field. A strong pre-mating barrier within these two forms has been demonstrated, as in a rice growing area from Burkina Faso where M and S molecular forms were known to be sympatric, yet only the swarms of the M form had been collected (Diabaté et al. 2003). However, a few mixed swarms of the two forms were collected one year later in Soumousso where no mixed swarms had previously been reported (Diabaté et al. 2006). Interestingly, recent studies from Mali showed that the M and S forms occupy segregated swarm sites without any mixed form swarms being observed (Diabaté et al. 2009, 2010, 2011), whereas in Burkina Faso the latest studies conducted between 2005 and 2006 reported relatively high frequencies of mixed swarms both in Soumousso and Vallée Kou (Dabiré et al. 2013). Even though there is a lack of consensus on many points, these two molecular forms were recently formally named as distinct species belonging to the *Anopheles gambiae* complex and termed *Anopheles coluzzii* and *Anopheles gambiae s.s.* for the M and S molecular forms, respectively (Coetzee et al. 2013).

The goal of this study was to investigate the swarming behaviour of *An. gambiae* M and S molecular forms over a relatively long period, 2006 to 2009, investigating the dynamics of the two forms in two rural villages (VK7 and Soumousso) that contrast in habitat, aiming to assess temporal and seasonal variations between molecular forms. The long term aim was to gather basic information on swarming and mating behaviour in these major malarial species that could help to better inform the effective use of sterile males or genetically modified mosquitoes to

4

suppress or replace vector populations and to further elucidate the mechanisms by which these closely related entities became reproductively isolated..

2. Materials and methods

2.1. *Sampling sites*

Vallée de Kou (11°24'N; 04°24'W) is about 30 km North-West of the second largest town of Burkina Faso, Bobo-Dioulasso, in the valley of the Kou River, a region where extensive rice cultivation was developed and has been practised since the 1970s. This area is organised into 7 villages, covering 7,200 ha surrounded by wooded savannah. As the Kou River flows all year round, it offers a permanent source of water for irrigation, hence allowing two crops of rice to be grown per year (July-November and January-May). Because of irrigation, the rice fields are highly productive permanent mosquito breeding sites. Additional anopheline breeding sites are depressions and rain puddles. Both M and S molecular forms of *An. gambiae* have been recorded at high densities during the rainy season, especially of the M-form: *c.* 200 bites person^{-1} night^{-1} were reported by Baldet et al. (2003). Also in 2003, swarms of *An. gambiae* M form were observed and the ecological parameters in the village 7 (VK7) were described (Diabaté et al. 2003). *Anopheles arabiensis* and *Anopheles funestus* were rare in this site.

Soumousso (11°00'46"N, 4°02'45"W) is a typical Guinean savannah village situated about 55 km East of Bobo-Dioulasso. There are two distinct seasons over the year, with an annual average rainfall of 1,000-1,200 mm, occurring between May and October. The breeding sites in the village consist mostly of rain-filled puddles and a semi-permanent swamp suitable for anopheline development. Three main malaria vectors are found, including both M and S molecular forms of *An. gambiae*, *An. funestus* and *An. nili*. *Anopheles arabiensis* is occasionally reported at low frequencies (<5% of *An. gambiae* s.l. samples). Since 2003 we have studied

swarming behaviour in the *An. gambiae* complex in Soumousso as M and S molecular form live in sympatry in this area. Mixed-swarms were first collected during a survey conducted in the rainy season of 2004 (Diabaté et al. 2006).

2.2. *Identification of physical markers associated with swarm formation*

All swarm sites were characterised each year by identifying the visual markers above which swarms were observed. These markers included physical objects such as woodpiles, piles of manure, patches of bare soil and other markers acting as a cue, attracting aggregation of male mosquitoes preceding the mating events. Each marker was described in terms of location and the nature of the object acting as swarm marker was recorded. Each swarm site was sampled during four consecutive days per month from July to October, corresponding to the rainy season when when mosquito populations are highest. The main swarms in VK7 and Soumousso were geo-referenced rigorously every month at the same places from 2006 to 2009, using the global positioning system (GPS) GARMIN, series GPSMAP® 62.

2.3. *Dynamic of swarming and mating behaviour*

Swarming and mating behaviour of M and S molecular forms of *An. gambiae* were described first in 2006 in both study sites and observations were repeated in subsequent years; the times of swarm initiation and dispersal were recorded and the height of the centre of each swarm was measured as its centre using a graduated wooden stick placed on the ground under the centre of the swarm. In 2007 we attempted to evaluate the dynamics of mating during several swarming events involving 20 individual swarms in VK7; the number of pairs leaving the swarm *in copula* and the swarm size, as judged by the number of males still present at the end of the swarming period, were recorded. A camera (Casio® EX-Z600 Camera, 6.0. Mega Pixels) fitted with a flashbulb was used around the expected end of swarming behaviour and pictures were analysed

6

by eye; single mosquitoes are clearly visible and the swarm size was estimated by counting the number of males using image J software J 1.45 (Wayne Rasband, National Institutes of Health, U.S.A.,*fr.wikipedia.org/wiki/ImageJ*). Then the number of pairs that were recorded leaving *in copula* was compared to the swarm size to investigate any correlation.

2.4.   Mosquito collections from swarms and indoor resting collections

Individual mosquitoes and pairs collected *in copula* were sampled from swarms in both sites using insect sweep-nets following a standard procedure (Dabiré et al., 2013) from June to October each year, except in 2006, when sampling was conducted from July to November. Indoor resting females were collected by spraying 10 randomly selected village houses near identified swarming sites with non-persistent aerosol insecticides (pyrethrum) on four consecutive days per month on the same days that swarm characteristics were collected. Ffemale mosquitoes were knocked-down and immediately retrieved from white sheets laid on the floor. Abdomens of all single and paired females collected within the swarms and those collected indoors were stored in alcohol at -20°C for later observation of insemination status and identification of the molecular form of the female and any spermatozoa present by PCR. An average of 60 male mosquitoes per swarm was sub-sampled from at least five swarms randomly selected each month and analysed by PCR to determine their molecular form.

To evaluate the insemination status of females collected within the swarms (females collected *in copula* and other females caught alone) we dissected spermathecae to visually check for the presence of spermatozoa. The male and female in each copula and the sperm dissected from within each female were identified by PCR to molecular form (Dabiré et al. 2013).

2.5. DNA extraction and PCR amplification

All mosquitoes caught *in copula* and those dissected for insemination status were systematically identified to species by PCR. Genomic DNA was extracted from whole individuals or legs according to the procedure described by Collins et al. (1987). Molecular forms were determined using the standard PCR-based diagnostic tool described by Favia et al. (2001). The few specimens from which a positive PCR result was not obtained (thought to be *An. arabiensis*) were re-tested to determine species using the PCR technique according to Scott et al. (1993).

2.6. Data analysis

The proportions of *An. gambiae* M and S molecular forms collected from mixed swarms and from indoor resting collections were compared by season and by year in each of the two sites. The frequency of mixed swarms was calculated according to the number of swarms and specimens successfully analysed by PCR. All comparisons were performed using Chi square or Anova tests.

3. Results

3.1. Location of swarming places in study sites

In the two sites, all swarms observed and collected from 2006 to 2009 were located inside or outside human households but always close to the village centre without any aggregation around the breeding sites, as previously thought. No swarms were reported more than fifty meters from the village. No particular segregation was observed between swarms of the two molecular forms at VK7 (Fig. 1A-D) and Soumousso (Fig. 1E-H), respectively. Swarms sites are defined as physical or visually conspicuous markers above which swarms occurred.

For simplicity, throughout the rest of the manuscript, a swarm observed on a particular evening will be referred to as being an 'M-form', 'S-form' or 'mixed-form swarm', depending on the

nature of the sub-sample of males randomly selected for PCR analysis from the sweep net sample collected from a swarm that evening. It is worth pointing out, however, that the sampling does not necessarily provide an accurate reflection of the total swarm composition on a given evening, particularly for those described as 'single-form' swarms; the other form may well have been present, but at low frequencies or before/after the sample was taken.

3.2. Swarm markers and swarm characteristics

Swarm sites in VK7 were mainly located outside (74.2%) but near human habitations, and the rest were recorded indoors (25.8%). The markers above which swarms formed were of 13 types: patches of bare ground, shelters, wood piles, manure piles, walls, wells, gallery, balcony, herbs, waste piles, indoor cooking fireplace, hen houses and large stones. The most frequently used markers were woodpiles, bare ground, shelters, wells and manure piles (Fig. 2A). The visual markers used by swarms in Soumousso (62% outside vs 38% indoors) did not differ in any obvious or consistent way from those observed in VK7, and the most commonly used markers were patches of bare ground, indoor sites (comprising bare soil), wells and maize fields (Fig. 2B).

Swarms of *An. gambiae* M form in VK7 were observed over all 13 marker types; swarms of the S molecular form were only found above manure piles, waste piles and galleries, though only 5 swarms of the S form were observed in total. Similarly, in Soumousso both molecular forms were found in swarms, dispersed around the site and sometimes occurring in the same areas. Some swarm markers, such as wells and patches of bare ground, were shared by the two molecular forms and also by *An. arabiensis*.

The shortest distance recorded between distinct swarms in VK7 was 3 m, and some satellite swarms (defined as "unstable" and relatively small swarms that formed after the principal swarms had formed, generally not appearing on subsequent days) formed close to an associated

principal swarm, at an average distance of 2 meters (Fig. 1A-D). Similar patterns were observed in Soumousso even though the number of swarms was lower (Fig. 1E-H).

The mean height of swarms measured in VK7 was 2m, compared to an average of 1.5m in Soumousso; the highest swarms were nearly 5m and the lowest around 0.5m above ground (Fig. 3). In VK7 the height of swarms varied between 1.5 m and 3m (mean ± se: 1.93 m±0.06) and sometimes higher when the swarm was disturbed (Fig. 3). Few swarms were observed at a height lower than 1 m, especially in Soumousso where the mean height was 1.42m±0.06, significantly lower than that in VK7 (t-test, t=6.34, *P*<0.05). No difference was observed between years , so only the mean height values are given to compare between sites.

3.3. Mating behaviour within An. gambiae swarms

The mating behaviour of the *An. gambiae* species complex was described first in 2006, describing the dynamics of mating in-depth in relation to the size of swarms (Dabiré et al. 2013) and followed up by a study from 2006-2009 (Sawadogo et al., 2013). Initiation of swarming in the M form at the VK7 study site occurred at a mean of 6.76±0.256 min after sunset, whereas S form swarms in Soumousso appeared significantly earlier, at 3.41±0.256 after sunset (Welch's Two sample t-test; t = 9.2515, df = 979.98, p < 2.2 x 10^{-16}).

The maximum number of mating pairs was observed when the swarm size reached 800-1,000 males, after which point the number of mating pairs decreased with increase in swarm size (Fig. 4).

3.4. Seasonal and annual variation of frequency of mixed swarms

Overall, 603 swarms were observed between June 2006 and October 2009 at VK7 and 372 in Soumousso. Because of the height of some swarms and the sampling method used, swarms reaching heights above 3.5m were not properly sampled and so were excluded from further analysis. Therefore, 250 swarms from VK7 and 205 from Soumousso were effectively sampled. Subsamples of 195 swarms from VK7 and all swarms collected from Soumousso were identified by PCR for their species and molecular form.

In VK7 swarms were mostly composed of pure swarms of M molecular form (175/195 tested, 89%), with only 4.6% (9/195) comprising the S form and 5.6% (11/195) of swarms being of mixed molecular form (Table 1). The mixed swarms were collected each year toward the end of the rainy season at varying frequencies: 11.3% of all swarms were of mixed molecular form in 2006, 3.3% in 2007, 5.0% in 2008 and 2.5% in 2009 (Fig. 5A). The relatively large proportion of pure swarms of M form in VK7 was associated with a predominance of M form in other samples, e.g. resting site catches (Fig. 5B), although the S form increased in proportion towards the end of the rainy season, peaking maximally in October (5-17%).

In Soumousso, by contrast, more mixed swarms were observed (24/205 tested, 12%) relative to the number of pure swarms of both S-form (147/205, 72%) and M form (27/205, 13%). In 2006 many more mixed swarms were found, at frequencies ranging from 75% to 12.5% of all swarms sampled (Table 2). The proportion of mixed swarms did not differ from 2006 to 2008 ($\chi^2 = 1.63$, ddl = 2, $P = 0.44$), but in 2009 the proportion decreased significantly ($\chi^2 = 21.71$, ddl = 3, $P < 0.05$). The proportion of the M and S forms within the mixed swarms showed that the number of M molecular form males was significantly lower than the S form, irrespective of the sampling month ($\chi^2=178$; df=2; $P<0.001$) (Fig. 5C). The proportion of mixed swarms varied from one year to another, although mixed swarms were mostly observed early in the rainy season, from June to July, or towards the end of the rainy season when the M form males increased slightly in density (Fig. 5D). Some individuals of *An. arabiensis* were also reported

within swarms of the M or S form in 2007 and 2009, particularly in June and July (Table 2). As in VK7, 60 satellite swarms were observed which did not appear regularly each month but were more frequent in August and September when mosquito densities were at their highest.

3.4.Molecular identification of pairing males and females caught in copula collected within the swarms and their sperm

In VK7 491 pairs were caught *in copula* and 100 single females were caught from 250 swarms, and in Soumousso 108 pairs were caught *in copula* (Tables 3 and 4) and 25 single females were also sampled from 60 swarms between July 2006 and October 2009.

In VK7 the mating pairs mostly comprised matched couples where both individuals were of the M-form (441 couples, 89.5%); 9.5% (47/491) of couples were made up only of the S form (Table 3). The proportion of mating pairs of mixed molecular form was low: 0.8% of all matings, including those collected from mixed swarms, comprising 3 mixed pairs caught in 2006 (2.4% of all pairs caught in this year) and one in 2008 (0.6% of pairs caught).

Except in 2006, all mating pairs caught in Soumousso, even those collected in mixed swarms, were composed of individuals of the same molecular form, most commonly the S form (107 females and respective males were of the S forms *vs* 0 of the M form) (Table 4). Only one mixed pairing was collected (1/108) corresponding to 4% (1/25) of pairs collected in 2006 being mixed-form. During the other three years of sampling no mixed pairs were caught.

Overall, 51 single females collected from swarms in VK7 (4, 16 and 30 in 2006, 2007 and 2009, respectively) and a further 8 collected in Soumousso in 2008 were analysed to determine the identity of their sperm.

All the inseminated females from both sites, including those caught *in copula* and single females caught within swarms, were found to contain sperm from a male of the same molecular

form as herself (Table 5). Identification of the molecular form of sperm collected from the four

mixed-form pairs revealed that even in these females the sperm was of the same molecular form

as themselves, indicating that the female had already mated with a male of the same molecular

form.

3.5.Frequency of M and S molecular forms from indoor collections

In VK7, significantly more M form than S form females were collected during indoor

collections (Fig. 5B), regardless of the month or the year (χ^2=152, df=2, *P*<0.001). However,

the proportion of S form females caught peaked in September 2006, reaching 15% of all

individuals caught.

Across the three years of surveillance, the M form dominated the site of VK7, with the

occurrence of the S form increasing toward the end of the rainy season. As in VK7, the

frequency of the two molecular forms, and also *An. arabiensis* which was also collected in

Soumousso indoor catches, reflected that observed in swarms, being significantly dominated

by the S form throughout the three year sampling period (χ^2=42.3, df=2, *P*<0.001) (Fig. 5D).

An. arabiensis was collected at frequencies less than 10% of individuals collected, especially

at the beginning of each rainy season, regardless the year. No hybrids were reported from any

of the mosquitoes collected indoors from either of the two sites.

4. Discussion

The investigation of swarming and mating behaviour in natural populations of *An. gambiae* s.s.

in Burkina Faso began in 2003 when swarms of the M molecular form were reported in VK7

and the authors found no mixed swarms (Diabaté et al, 2003). These findings suggested the

existence at that time of a strong premating reproductive barrier based on spatial or temporal

segregation between the two forms, preventing hybridisation. That supported the results of

Charwood et al. (2002), who did not observe any mixed swarms. But for the first time, in 2004, a few mixed swarms, largely dominated by the S form with a few individuals of the M form, were reported in Soumousso, a village in a wooded savannah region. The low frequency of mixed swarms observed did not match the expected frequency based on the much larger proportions of M form collected indoors, which sometimes reached 18% (Diabaté et al. 2006). Through extended investigations from July 2006 to October 2009, we observed a great number of mixed swarms, at frequencies that varied temporally in the two sites.

Swarms were always initiated by a male mosquito (termed the swarm "precursor"), which arrived at the swarm site and flew upwards in a zigzag pattern. It was then joined by more males 5-10 minutes later. After the first 10 minutes the first couples could be observed leaving the swarm *in copula.* The peak of couple formation occurred 15 - 25 minutes after the first pair formed. The total duration of swarming events varied with season, the swarming period being relatively longer in July than in October. In addition, the density of males within the swarm and the number of mating pairs were highest in August-September, corresponding to the middle of the rainy season. The number of mosquitoes observed taking part in swarms, recorded as the number of males present in the swarm at the peak of the swarming period, varied with the season, with the biggest swarms being consistently occurring in August-September, which correlated with the density of resting mosquitoes collected indoors. These two months are the most favourable period for mosquito reproduction, in that breeding sites are most prevalent and in a favourable condition for mosquito development. Although this information was supported by field observation, more detailed observations would be needed to back it up with data. Interestingly, the initiation of *An. gambiae* swarms could be predicted by the appearance, about five minutes before the first male mosquito, of one or two dragonflies (Odonata) flying around the swarm site at the same height as the swarm later formed. The dynamic of mosquito mating within swarms could vary when one was disturbed by predatory activities. Indeed, sometimes

more than ten dragonflies entered a swarm and hunted mosquitoes, which often led to a decrease in the number of couples forming. After sunset, bats were observed to enter swarms and catch mosquitoes in flight. It was difficult to investigate the scale of their predation, but this was also likely to affect mating behaviour and overall mating success within a swarm.

Observations of 20 individual swarms performed in 2007 in VK7 indicated that the number of pairings increased with the size of the swarms up to a maximum optimal size; this number seemed to be underestimated when the swarm size was up to 1,000 mosquitoes. Direct observation of swarms in the field noted that when the size of a swarm reaches a maximum, smaller satellite swarms formed nearby, in which mating could be observed, and may have reduced the number of mating pairing in the principal swarms. It is likely that the number of mating pairs may be reduced or mating interrupted by predatory activities (intrusion of dragonflies or bats).

The high frequencies of mixed swarms and the occurrence of mixed pairings would suggest there was the potential for hybrid matings to occur, although from a total of 300 inseminated females examined across four years' collections, no hybrid inseminations were reported (although the sample size was relatively low). The key finding of our results is that, although numerous pairs were caught *in copula* from mixed swarms, the great majority were composed of males and females from the same molecular form, in each of the four years of surveillance. We collected a few pairs apparently *in copula*, but analysis of the sperm carried by the female confirmed that the sperm stored by the females in the spermathecae were of the same molecular form as herself (Dabiré et al. 2013). This is evidence for strong assortative mating between forms and is in agreement with the low frequency of M/S hybrids or between-form insemination rates reported previously (Tripet et al. 2003). It suggests also that the commonly reported low level of cross pairings between forms could be the result of assortative mate choice rather than

15

the outcome of natural selection against hybrids in the field, i.e., pre-zygotic barriers to hybridisation.

No hybrid M/S form individuals were reported from indoor samples or mating couples from either of the two sites, indicating that the occurrence of mixed swarms and/or mixed pairings does not necessarily lead to cross-mating and hybridization (Dabiré et al, 2013). Again, we are aware that the sample sizes of indoor collected mosquitoes were too low to be certain that no hybrids were present. The relative frequencies of mixed swarms and the proportion of each form within the mixed swarms were as expected from indoor resting collections. Furthermore, the analysis of indoor collections of adult mosquitoes within the *An. gambiae* group is not the most precise way to investigate the presence of hybrids, as reflected by the discovery of Riehle et al. (2011) of a cryptic group of *An. gambiae* from larval collections indicating the probable occurrence of hybridisation. Further, the spermatozoa from the five mixed pairs were identified to be the same molecular form as the female, and not the male. If these five mixed pairings were able to achieve a successful mating, assuring hybridisation between the two forms, their rate should probably match the low frequency of cross mating of 1.4% already observed by Tripet et al. (2001, 2003).

Our findings have confirmed the existence of strong assortative mating between the M and S molecular forms of *An. gambiae* in the field as proposed by Dabiré et al. (2013) in Burkina Faso and as reported in Mali (Tripet et al. 2005; Diabaté et al. 2003, 2006, 2009 and 2011). The mechanism of this assortative mating seems to be temporal and seasonal segregation, so that each form dominates at different times. During the present study, no spatial segregation was observed in the two villages between the two forms across the four years of surveillance. Similarly, the use of visual markers as swarm sites were not specifically associated to one molecular form in these areas, although the two forms appeared to favour visually different typs of markers. Physical markers were mostly used by the M form in VK7, whilst the S form

swarmed more frequently above open areas in Soumousso. This was not the case in Mali, where Diabaté et al. (2009, 2011) described distinctive swarming systems characterised by segregation between M and S forms in terms of space and visual markers, leading to mostly assortative mating.

The occurrence of mixed swarms and mating pairs indicates that each molecular form is able to exploit many available visual swarming cues without any formal preference for specific markers. The sharing of the main swarm markers would facilitate the overlapping of swarms of different types, increasing the frequencies of mixed swarms. It is likely, therefore, that segregation is achieved either by the fact that only females of the predominant form in the mixed swarm approach the swarm, or that individuals within a swarm are able to select mates of the same molecular type. From the current data it is not possible to establish whether it is mate recognition or swarm recognition by the females, which is leading to segregation. However in the few pairings that were observed between males and females of different molecular types, no cross insemination was recorded, suggesting the existence of intrinsic unknown recognition mechanisms reducing hybridisation between the two forms.

The mechanisms enabling recognition of mates of the same molecular form could include the detection of volatile pheromones and/or hydrocarbons or other secreted proteins. This hypothesis is supported by a study of Yuval et al. (2006) that associates contact pheromones with the sexual behavior of several biting flies. But these mechanisms are not clearly elucidated, and Tripet et al. (2005) failed to identify the role of sex proteins in *An. gambiae* assortative mating. Nevertheless, the authors suggested that volatile stimuli like pheromones might be the cues forming the basis of pre-mating isolation. Wing beat frequencies may also play a key role in form recognition, as demonstrated by the recent work of Pennetier et al. (2010).

Our data provide a basis for evaluating the mating dynamics of *An. gambiae* through recording > 20 natural swarms. The number of pairs leaving swarms *in copula* appears to be density-dependant, increasing with male density within the swarm up to an optimal level, but decreasing when the swarm becomes bigger than that. Our sampling protocol, which was based on sweep netting, was not perfect and may not allow us to record with accuracy all pairs leaving the swarm *in copula*, especially when the swarm size was large (swarms reached sizes of up to 1,000 males). Furthermore, the time at which swarms initially formed did not vary significantly from one year to the next in the same sites. There is an indication that S form swarms appear earlier in Soumousso than M form swarms form in VK7,. The swarming height was also significantly different between the forms; S swarms in Soumousso appeared at significantly lower heights than the M form in VK7. However, swarming height was dynamic and varied when the swarm was disturbed by the activities of predators' activities or according to changes in the weather.

Our results show that the swarming and mating systems of the *An. gambiae* M and S molecular forms (now known as *An. coluzzii* and *An. gambiae s.s.*, respectively) are very stable, as the data were quite similar from one year to the next, indicating assortative mating between the two species. The time of swarm formation seemed to be conserved, without major variation between years and sites. This information is essential for the successful implementation of the sterile insect technique or the alternative use of genetically modified mosquitoes for reduction or suppression of this important malaria vector.

Conflict of interests

The authors declare no conflicts of interest.

Authors' contributions

SPS conducted fieldwork and contributed to data analyses. TKH, MH, MN and OKR contributed to fieldwork. JR drew the maps of swarms. TB, LCG, KP, SF, CC, GG, DA, LRS and GJ contributed to drafting and revising the manuscript. DKR designed the study, contributed to the data analyses and wrote the manuscript.

References

Baldet, T., Diabaté, A., Guiguemdé, T.R., 2003. Etude de la transmission du paludisme en 1999 dans la zone rizicole de la Vallée du Kou (Bama) Burkina Faso. Cahiers Santé 13, 56-60.

Caputo, B., Nwakanma, D., Jawara, M., Adiamoh, M., Dia, I., Konate, L., Petrarca, V., Conway, D.J., della Torre, A., 2008. *Anopheles gambiae* complex along The Gambia river, with particular reference to the molecular forms of *An. gambiae s.s.*. Malar. J. **7**, 182.

Caputo, B., Santolamazza, F., Vicente, J.L., Nwakanma, D.C., Jawara, M., Palsson, K., Jaenson, T., White, B.J., Mancini, E., Petrarca, V., Conway, D.J., Besansky, N.J., Pinto, J., della Torre, A., 2011. The "Far-West" of *Anopheles gambiae* Molecular Forms. PLoS One 6, e16415.

Charlwood, J.D., Jones, M.D.R., 1979. Mating in the mosquito *Anopheles gambiae* s.l. I. Close range and contact behaviour. Physiol. Entomol 4, 111-120.

Coetzee, M., Hunt, R., Wilkerson, R., Della Torre, A., Coulibaly, B.M., Besansky, N., 2013. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. Zootaxa 3619, 246-274.

Collins, F.H., Mendez, M.A., Rasmussen, M.O., Mehaffey, P.C., Besansky, N.J., Finnerty, V., 1987. A ribosomal RNA gene probe differentiates member species of the *Anopheles gambiae* complex. Am J Trop Med Hyg 37, 37-41.

Coluzzi, M., Petrarca, V., Di Deco, M.A., 1985. Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae*. Bolletino di Zoologia 52, 45-63.

Coluzzi, M., Sabatini, A., della Torre, A., Di Deco, A., Petrarca, V., 2002. A polytene chromosome analysis of the *Anopheles gambiae* complex. Science 298, 1415-1418.

Cooter, R.J., 1989. Swarm flight behavior in flies and locusts, in: Goldsworthy, G.J., Wheeler C.H. (Eds) Insect flight. CRC press, Boca Raton Florida, pp. 165-203.

Dabiré, K.R., Sawadogo, P.S, Diabaté, A., Toé, K.H, Kengne, P., Ouari, A., Costantini, C., Gouagna, L.C., Simard, F., Baldet, T., Lehmann, T., Gibson, G., 2013. Assortative mating in mixed swarms of the mosquito *Anopheles gambiae* s.s. M and S molecular forms, in Burkina Faso, West Africa. Med. Vet. Entomol. doi: 10.1111/j.1365-2915.2012.01049.x.

Della Torre, A., Fanello, C., Akogbeto, M., Dossou-Yovo, J., Favia, G., Petrarca, V., Coluzzi, M., 2001. Molecular evidence of incipient speciation within *Anopheles gambiae* s.s. in west Africa. Insect Mol. Biol. 10, 9-18.

Della Torre, A., Costantini, C., Besansky, N.J., Caccone, A., Petrarca, V., Powell, J.R., Coluzzi, M., 2002. Speciation within *Anopheles gambiae* - the glass is half full. Science 298, 115-117.

Della Torre, A., Tu, Z., Petrarca, V., 2005. On the distribution and genetic differentiation of *Anopheles gambiae* s.s. molecular forms. Insect Biochem Mol Biol 35, 755-769.

Diabaté, A., Baldet, T., Brengues, C., Kengne, P., Dabiré, K.R., Simard, F., Chandre, F., Hougard, J.M., Hemingway, J., Ouédraogo, J.B., Fontenille, D., 2003. Natural swarming behaviour in the molecular M form of *Anopheles gambiae*. Trans. R. Soc. Trop. Med. Hyg. 97, 713-716.

Diabaté, A., Dabiré, K.R., Kengne, P., Baldet, T., Ouari, A., Costantini, C., Simard, F., Fontenille, D., 2006. Mixed-swarms of the molecular M and S forms of *Anopheles gambiae* in a sympatric area from Burkina Faso. J. Med. Entomol. 43, 480-483.

Diabaté, A., Dao, A., Yaro, A.S., Adamou, A., Gonzalez, R., Manoukis, N.C., Traoré, S.F., Gwadz, R.W., Lehmann, T., 2009. Spatial swarm segregation and reproductive isolation between the molecular forms of *Anopheles gambiae*. Proc. Biol. Sci. 276, 4215-4222.

Diabaté, A., Yaro, A.S., Dao, A., Diallo, M., Huestis, D.L., Lehman, T., 2011. Spatial distribution and male mating success of *Anopheles gambiae* swarms. BMC Evol. Biol. 11, 184.

Favia, G., Lanfrancotti, A., Spanos, L., Sidén-Kiamos, I., Louis, C., 2001. Molecular characterization of ribosomal DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae* ss. Insect Mol. Biol. 10, 19-23.

Lawniczak, M.K.N., Emrich, S.J., Holloway, A.K., Regier A.P., Olson, M., White, B., Redmond, S., Fulton, L, Appelbaum E., Godfrey J., Farmer C., Chinwalla A., Yang S.P., Minx P., Nelson, J., Kyung, K., Walenz, B.P, Garcia-Hernandez, E., Aguiar, M., Viswanathan, L.D., Rogers, Y.H., Strausberg, R.L., Saski, C.A., Lawson, D., Collins, F.H., Kafatos, F.C., Christophides, G.K., Clifton, S.W., Kirkness, E.F., Besansky, N.J., 2010. Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. Science 330, 512-514.

Oliveira, E., Salgueiro, P., Palsson, K., Vicente, J.L., Arez, A.P., Jaenson, T.G., Caccone, A., Pinto, J., 2008. High levels of hybridization between molecular forms of *Anopheles gambiae* from Guinea Bissau. J. Med. Entomol. 45, 1057-1063.

Pennetier, C., Warren, B., Dabiré, R., Russell, I.J., Gibson, G., 2010. ''Singing on the wing'' as a mechanism for species recognition in the malarial mosquito *Anopheles gambiae*. Curr. Biol. 20, 131-136.

Persiani, A., Di Deco, M.A., Petrangeli, G., 1986. Osserzioni di laboratorio su polimorfismi da inversione originati da incroci tra popolazioni diverse di *Anopheles gambiae* s.s.. Annal de la Istito Superio de la Sanita 22, 221-224.

Riehle, M.M., Guelbeogo, W.M., Gneme, A., Eiglmeier, K., Holm, I., Bischoff, E., Garnier, T., Snyder, G.M., Li, X., Markianos, K., Sagnon, N., Vernick, K.D., 2011. A cryptic

subgroup of *Anopheles gambiae* is highly susceptible to human malaria parasites. Science 331, 596-598.

Sawadogo S.P., Costantini C., Pennetier C., Diabaté A., Gibson G., Dabiré K.R., 2013. Influence of circadian rhythms and environmental factors on the timing of natural swarming behaviour in *Anopheles gambiae* M and S molecular forms in Burkina Faso, West Africa. Parasit. Vectors (in press).

Scott, J.A., Brogdon, W.G., Collins, F.H., 1993. Identification of a single specimen of *An. gambiae* complex by polymerase chain reaction. Am. J. Trop. Med. Hyg. 49, 520-529.

Sullivan, R.T., 1981. Insect swarming and mating. Florida Entomologist 64, 44-65.

Touré, Y.T., Petrarca, V., Traoré, S.F., Coulibaly, A., Maïga, H.M., Sankaré, O., Sow, M., Di Deco, M.A. & Coluzzi, M. (1998) The distribution and inversion polymorphism of chromosomally recognised taxa of the *Anopheles gambiae* complex in Mali, West Africa. Parasitologia 40, 477-511.

Tripet, F., Touré, Y.T., Taylor, C.E., Norris, D.E., Dolo, G., Lanzaro, G.C., 2001. DNA analysis transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. Mol. Ecol. 10, 1725-1732.

Tripet, F., Touré, Y.T., Guimogo D., Lanzaro, G.C., 2003. Frequency of multiple inseminations in field-collected *Anopheles gambiae* revealed by DNA analysis of transferred sperm. J. Med. Entomol. 68, 1-5.

Tripet, F., Thiemann, T., Lanzaro, G.C., 2005. Effect of seminal fluids in mating between M and S forms of *Anopheles gambiae*. J. Med. Entomol. 42, 596-603.

Wondji, C., Simard, F., Fontenille, D., 2002. Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of *Anopheles gambiae* in an area of sympatry. Insect Molecular Biology 11, 11-19.

Yuval, B., 2006. Mating systems of blood-feeding flies. Annu. Rev. Entomol. 51, 413-440.

Figure Legends:

Fig. 1. Characteristics and location of swarming sites within Vallée de Kou (VK7) and Soumousso, south-western Burkina Faso. Land use and location of family compounds in VK7 (panels A-D) and Soumousso (E-H). The distribution of larval breeding sites and sampled swarms (M or S form or mixed-form) is shown in 2006 (A and E), 2007 (B and F), 2008 (C and G) and 2009 (D and H).

Fig. 2. Swarm markers used by *Anopheles gambiae* and *Anopheles arabiensis* and their relative frequency among swarms observed in A) Vallée du Kou (VK7) and B) Soumousso.

Fig. 3. Variation of swarm height, measured as distance between the ground and the approximate centre of the swarm, in A) Vallée du Kou (VK7) and B) Soumousso. Each data point represents a single observed swarm, and the lines represent the average swarm height in each study site.

Fig. 4. Relationship between swarm size and number of mating couple observed in swarms of *Anopheles gambiae* M form in Vallée du Kou (VK7) (regression test, $r^2= 0,753$).

Fig. 5. Proportion of M and S forms of *Anopheles gambiae* and *Anopheles arabiensis* males sampled from swarms (A) and females collected indoors (B) in Vallée du Kou (VK7) and respectively in Soumousso (C-D).

Table Legends:

Table 1

Seasonal and annual dynamics of the frequency of M and S molecular forms of *Anopheles gambiae* and *Anopheles arabiensis* in segregated and mixed-form swarms in A) Vallée du Kou (VK7) and B) Soumousso. Nb = number, * = *Anopheles arabiensis*.

Table 2

Seasonal and annual dynamics of M and S molecular forms of *Anopheles gambiae* among swarms: frequencies of pure form swarms, mixed swarms (Nb = number, * *Anopheles arabiensis*)

Table 3

Molecular identification of the *An. gambiae* molecular form of individuals within pairs of mosquitoes caught *in copula* in Vallée du Kou (VK7) from 2006 to 2009. M = M-form swarm, S = S-form swarm, X = mixed-from swarm.

Table 4

Molecular identification of the *An. gambiae* molecular form of individuals within pairs of mosquitoes caught *in copula* in Soumousso from 2006 to 2009. M = M-form swarm, S = S form swarm, X = mixed-from swarm.

Table 5

Molecular identification of pairs of *An. gambiae* mosquitoes caught *in copula* and the spermatozoa collected from the spermathecae of the related females in Soumousso and Vallée du Kou (VK7) between 2006 and 2009.

Table 1: Seasonal and annual dynamics of the molecular forms among swarms: frequencies of pure form swarms, and mixed swarms in VK7 (Nb=number)

Sampling site VK7	Nb of swarms sampled	Swarm points ID	Nb of swarms tested	Pure M-form	Pure S-form	Mixed swarm M/S	% mixed
2006							
July	16	sw1, sw2, sw3, sw4, sw5, sw6, sw7, sw8, sw9,sw10	16	16	0	0	0
August	15	sw1, sw3, sw4, sw5, sw7, sw7', sw8, sw8', sw13, sw13', sw14	15	15	0	0	0
September	15	sw3, sw4, sw4', sw4", sw5, sw7, sw8, sw8", sw11, sw12, sw13, sw14	15	13	0	2	13.3
October	15	sw4, sw4', sw4", sw5, sw7, sw12, sw13, sw14	15	11	0	4	26.6
November	6	sw3", sw4, sw4", sw11, sw12, sw13	6	3	2	1	16.6
2007							
June	4	sw4, sw12''', sw13, sw14	4	4	0	0	0
July	4	sw4, sw12''', sw13, sw14	4	4	0	0	0
August	5	sw4, sw12''', sw13, sw14	5	5	0	0	0
September	6	sw4, sw4', sw4'b, sw4'', sw4'''b sw13	6	3	2	1	16.6
October	9	sw4, sw4b, sw4'b, sw4"", sw12, sw12", sw13, sw14, sw14b	9	8	1	0	0

2008

June	4	sw4', sw12, sw13, sw14	4	4	0	0	0
July	4	sw4', sw4'b, sw12''', sw13	4	4	0	0	0
August	4	sw4, sw11, sw12, sw4	4	4	0	0	0
September	5	sw2', sw4', sw11, sw13, sw14	5	4	0	1	25
October	4	sw4c, sw12b, sw12c, sw14'	4	2	2	0	0
2009							
June	15	sw2, sw2', sw4'', sw5, sw6, sw12, sw12'', sw12'''', sw13, sw14, sw16, sw16', sw16'', sw17, sw17''	15	15	0	0	0
July	25	sw1', sw3', sw4', sw4'b, sw4'', sw4''c, sw4'''', sw5, sw12', sw12'', sw12''', sw12'''', sw12''''''', sw13, sw14, sw14', sw14'''', sw14'''''', sw16, sw16'', sw16''', sw17, sw17', sw18, sw18'	15	15	0	0	0
August	38	sw1, sw2, sw3, sw4', sw4d, sw4''b, sw12, sw12b, sw12', sw12'', sw12''', sw13, sw14, sw14', sw14'''''', sw17, sw20, sw24, sw31, sw32, sw33, sw34, sw35, sw36	16	15	0	1	6.2
September	40	sw3', sw4', sw4'c, sw4'', sw4''c, sw4''d, sw4'''', sw12, sw12', sw12'', sw12''', sw12''''', sw12'''''', sw12'''''''', sw12''''''''b, sw13, sw14, sw14', sw14''', sw14'''''', sw16, sw16'', sw16''', sw17, sw17', sw20, sw31, sw32, sw35	17	15	2	0	0
October	16	w4', sw4'', sw4''', sw5, sw12b, sw12'', sw12''', sw12'''''', sw13, sw14', sw16'', sw20, sw31	16	15	0	1	6.2

sw1: principal swarm ID, sw1': satellite swarm ID

26

Table 2: Seasonal and annual dynamic of the molecular forms among swarms: frequencies of pure form swarms, mixed swarms (Nb=number, * *An. arabiensis*)

Sampling site **Soumousso**	Nb of swarms sampled	Swarm site ID	Nb of swarms tested	Pure M-form	Pure S-form	Pure arabiensis	Mixed swarm M/S	% mixed
2006								
July	8	sw2, sw7, Sw9, sw13, sw15, sw17, sw18, sw19	8	0	2	0	6	75
August	8	sw1, sw2, sw3, sw4, sw7, sw8, sw9, sw11	8	0	7	0	1	20
September	8	sw1, sw2, sw3, sw4, sw6, sw7, sw8, sw9	8	0	7	0	1	12.5
October	8	sw2, sw2", sw3, sw3' sw4, sw4', sw4", Sw5	8	0	6	0	2	40
2007								
June	6	sw1, sw1', sw2, sw2', sw3, sw4	6	2	3	0	0 (+1*)	0 (16.6)
July	17	sw1, sw1", sw2, sw2', sw3, sw4, sw4', sw5, sw5', sw6, sw7, sw7' Sw8, sw8', sw12	17	2	11	0	4	36.3
August	6	sw1", sw2, sw3, sw4, sw4"", sw7	6	1	4	0	1	16.6
September	7	sw1, sw1", sw2, sw2"", sw3, sw3', sw4	7	2	5	0	0	0
October	5	sw1, sw1', sw2, sw3, sw4	5	0	5	0	0	0
2008								
June	5	sw2, sw3, sw4', sw7, sw8	5	0	2	0	3	60
July	5	sw2, sw3, sw4, sw6, sw7	5	1	2	0	2	40

August	5	sw3, sw3", sw5, sw6, sw7	5	0	5	0	0	0
September	5	sw3, sw4, sw4'b, sw5, sw6	5	1	4	0	0	0
October	6	sw1", sw2b, sw3, sw3', sw4, sw5	6	0	4	0	2	33.3
2009								
June	16	sw3, sw4b, sw4'''', sw4''''', sw11, sw12, sw14, sw15, sw16, sw17, sw18, sw19, sw20, sw21, sw22	16	8	4	0	2 (+2*)	12.5(33.3)
July	24	sw1, sw1", sw1''', sw2, sw3, sw3", sw4'''', sw4''''', sw7', sw8, sw14, sw17, sw18, sw22, sw25, sw26, sw27	24	8	14	0	2	8.3
August	28	sw1, sw1", sw1''', sw2, sw3, sw3', sw4, sw4''''', sw5, sw5", sw10, sw11, sw14, sw20, sw24, sw25, sw27, sw28, sw29, sw30, sw31, sw32, sw33, sw34, sw35	28	2	24	0	0 (+2*)	0(7)
September	18	sw1, sw1', sw1", sw1''', sw2''''', sw3', sw3", sw5', sw11, sw28, sw31, sw36	18	0	18	0	0	0
October	20	sw1, sw1', sw1", sw2, sw2''''', sw3', sw3", sw3'''', sw3''''', sw5, sw11, sw20, sw25, sw27, sw28, sw30, sw31, sw33	20	0	20	0	0	0

sw1: principal swarm ID, sw1': satellite swarm ID

Table 3: Molecular identification of the pairs of mosquitoes caught *in copula* in VK7 from 2006 to 2009. M, M-form swarm; S, S-form swarm; X, mixed-from swarm.

Sampling site **VK7**	Nb of pairs collected	ID of Swarm sampled	Pure M-form	Pure S-form	Mixed pairs M/S	% mixed
2006						
July	4	sw3	4	0	0	0
August	26	sw3, sw5, sw7, sw14,	26	0	0	0
September	48	sw3, sw4, sw4', sw8, sw12, sw13, sw14	34	14	0	0
October	43	sw4, sw4', sw5, sw12, sw13, sw14	13	27	3	6.98
November	3	sw4	0	3	0	0
Total	*124*		*77*	*44*	*3*	*2.42*
2007						
June						
July						
August	29	sw4, sw13, sw14	29	0	0	0
September	46	sw4, sw4', sw4'b, sw4'', sw4'''b, sw13	45	1	0	0

October	59	sw4, sw4b, sw4'b, sw4"", sw12, sw12", sw13, sw14, sw14b	59	0	0	0
Total	*134*		*133*	*1*	*0*	*0*
2008						
June	20	sw4', sw12, sw13, sw14	20	0	0	0
July	26	sw4', sw4'b, sw12"', sw13	26	0	0	0
August	32	sw4, sw11, sw12, sw4	32	0	0	0
September	51	sw2', sw4', sw11, sw13, sw14	49	1	1	1.96
October	36	sw4c, sw12b, sw12c, sw14'	32	0	0	0
Total	*165*		*163*	*1*	*1*	*0.6*
2009						
June						
July	6	sw1', sw5	6	0	0	0
August	19	sw12"', sw13, sw14, sw14', sw14""', sw31, sw32, sw34, sw36	19	0	0	0
September	35	sw4", sw13, sw12"'	35	0	0	0
October	8	sw4', sw4", sw12"', sw14', sw16"	7	1	0	0
Total	*68*		*68*	*1*	*0*	*0*

Table 4: Molecular identification of the pairs of mosquitoes caught *in copula* in Soumousso from 2006 to 2009

Sampling site Soumousso	Nb of pairs collected	ID of swarm sampled	Pure M-form	Pure S-form	Mixed pairs M/S	% mixed
2006						
July						
August	17	sw1, sw2, sw3, sw11	0	17	0	0
September	6	sw2	0	5	1	16.67
October	2	sw2	0	2	0	0
Total	**25**			24	1	4
2007						
June						
July	1	sw1	0	1	0	0
August	1	sw1	0	1	0	0
September		-				
October		-				
Total	**2**	-		**2**	**0**	**0**

2008					
June		-			
July		-			
August	8	sw3, sw5, sw5'	0	8	0	0
September	36	sw3, sw4, sw4'b, sw5, sw6	0	36	0	0
October	1	sw1"	0	1	0	0
Total	**45**		**0**	**45**	**0**	**0**
2009					
June		-			
July	4	sw1, sw3, sw5'	0	4	0	0
August	18	sw2, sw3", sw5, sw11, sw28, sw31	0	18	0	0
September		-			
October	14	sw2"''', sw3', sw11, sw28, sw31	0	14	0	0
Total	**36**		**0**	**36**	**0**	**0**

Table 5: Molecular identification of the pairs of mosquitoes caught *in copula* in Soumousso and VK7 from 2006 to 2009

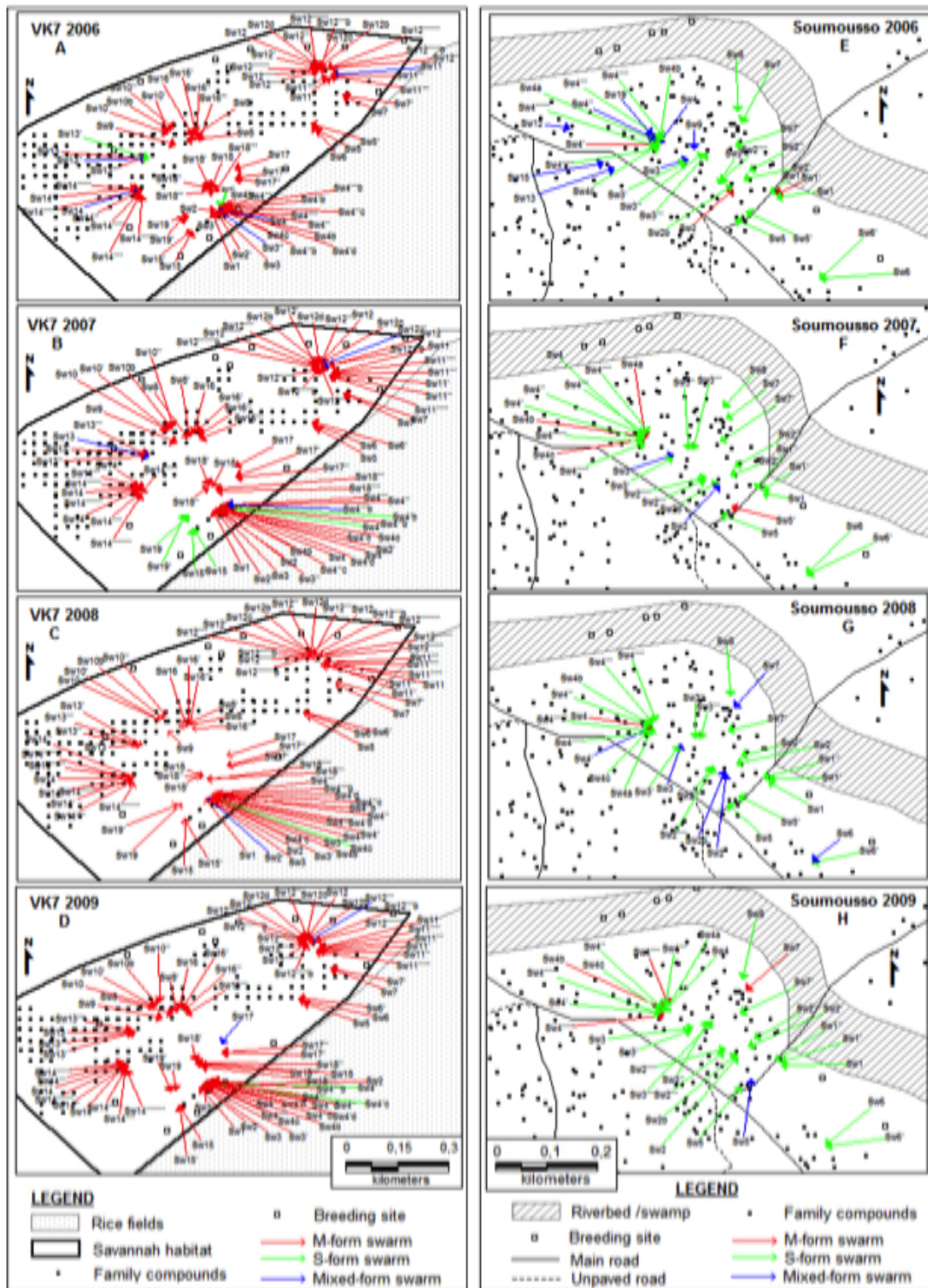| Locality/Date	Number of pairs dissected for their sperm	Females		Males		Sperm		Number of females solo tested	Females		Sperm	
		M	S	M	S	M	S		M	S	M	S
VK7												
2006	18	7	11	8	10	7	11	4	3	1	3	1
2007	75	74	1	74	1	74	1	16	16	0	16	0
2008	162	160	2	161	1	160	2	30	30	0	30	0
2009												
Total	**255**	**241**	**13**	**243**	**11**	**241**	**13**	**51**	**49**	**2**	**49**	**1**
Soumousso												
2006												
2007												
2008	45	0	45	0	45	0	45	8	0	8	0	8
2009												
Total	**45**	**0**	**45**	**0**	**45**	**0**	**45**	**8**	**0**	**8**	**0**	**8**

Figures

Fig. 1. Characteristics and location of swarming sites within Vallée de Kou (VK7) and Soumousso, south-western Burkina Faso. Land use and location of family compounds in VK7 (panels A-D) and Soumousso (E-H). The distribution of larval breeding sites and sampled swarms (M or S form or mixed-form) is shown in 2006 (A and E), 2007 (B and F), 2008 (C and G) and 2009 (D and H).
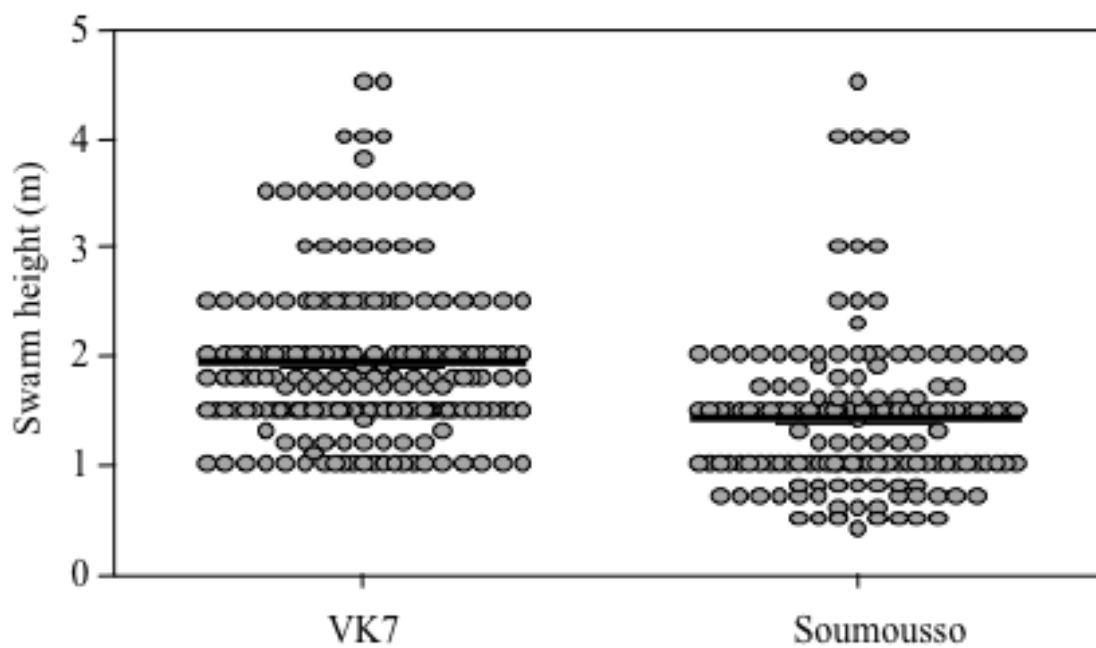
Fig. 2. Swarm markers used by *Anopheles gambiae* and *Anopheles arabiensis* and their relative frequency among swarms observed in A) Vallée du Kou (VK7) and B) Soumousso.

Fig. 3. Variation of swarm height, measured as distance between the ground and the approximate centre of the swarm, in A) Vallée du Kou (VK7) and B) Soumousso. Each data point represents a single observed swarm, and the lines represent the average swarm height in each study site.
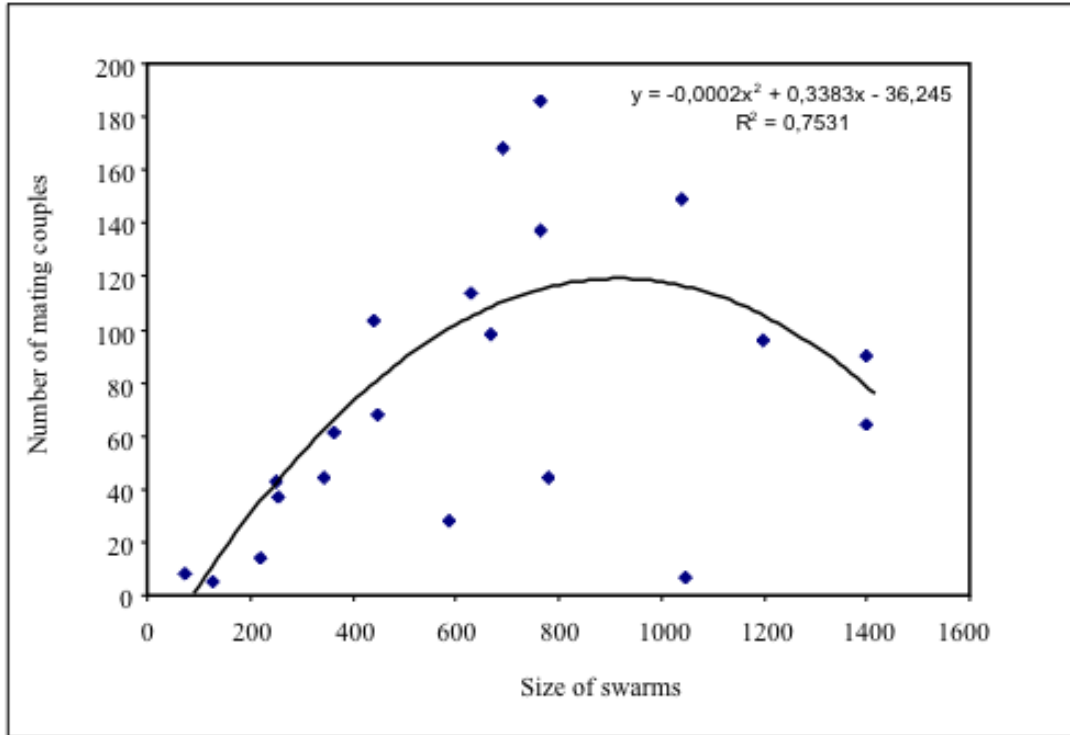
Fig. 4. Relationship between swarm size and number of mating couple observed in swarms of *Anopheles gambiae* M form in Vallée du Kou (VK7) (regression test, $r^2 = 0,753$).
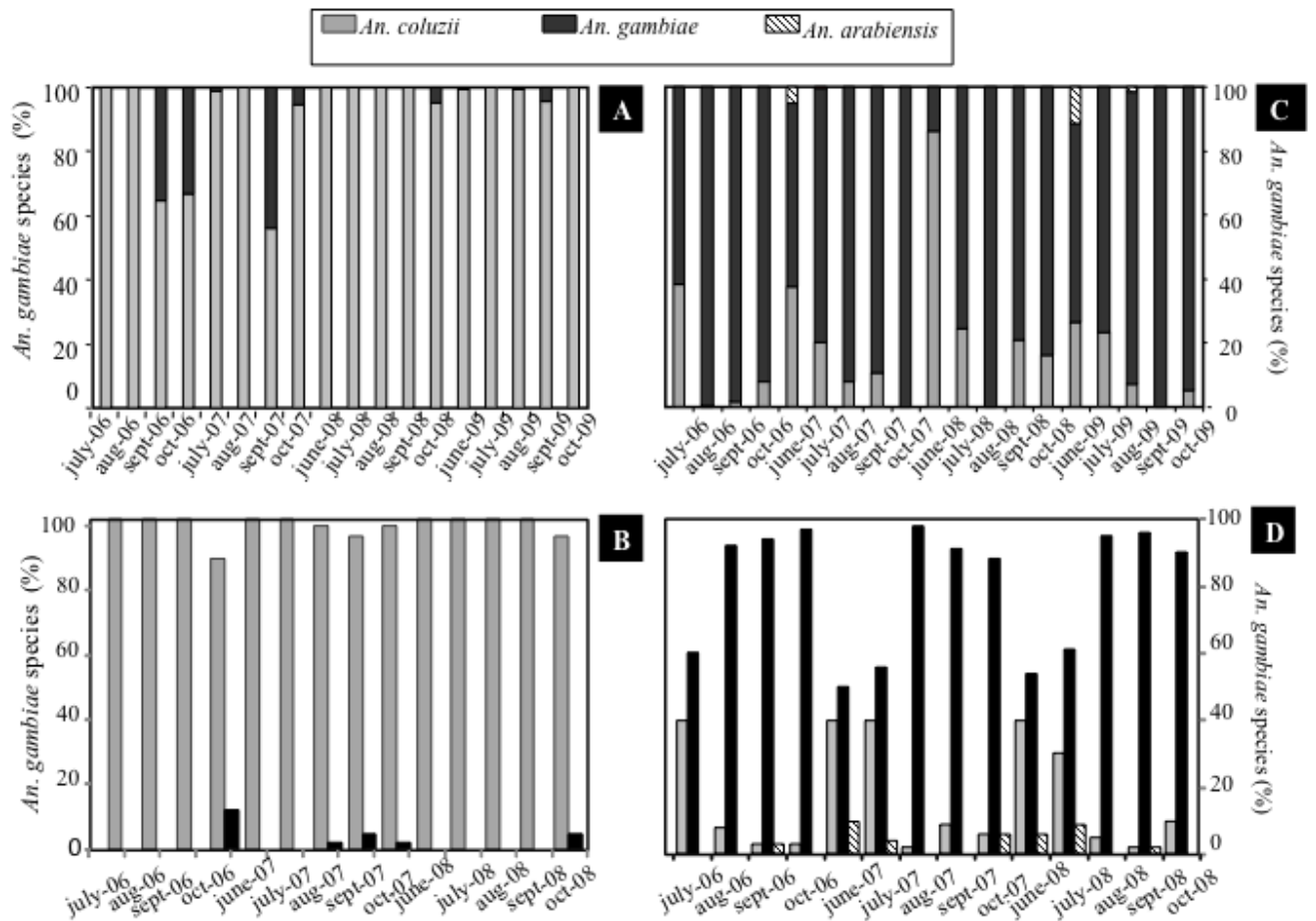
Fig. 5. Proportion of M and S forms of *Anopheles gambiae* and *Anopheles arabiensis* males sampled from swarms (A) and females collected indoors (B) in Vallée du Kou (VK7) and respectively in Soumousso (C-D).