We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter 6

Speciation in *Anopheles gambiae* — The Distribution of Genetic Polymorphism and Patterns of Reproductive Isolation Among Natural Populations

Gregory C. Lanzaro and Yoosook Lee

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/56232

1. Introduction

The African malaria vector, *Anopheles gambiae*, is characterized by multiple polymorphic chromosomal inversions and has become widely studied as a system for exploring models of ecological speciation. An attempt to develop a molecular diagnostic for the chromosomal forms of *A. gambiae* s.s. led to the development of a PCR-based diagnostic to differentiate M and S molecular forms based on a marker located on the X chromosome. Near complete reproductive isolation between M and S molecular forms has led to the suggestion that *A. gambiae* is in early stages of speciation. Comparative genomic studies have been applied to gain an understanding of the evolutionary process resulting in these forms, but models based on these studies currently lack consensus. Furthermore, various studies suggest further subdivisions within each molecular form. These topics are discussed and suggestions for further research needed to elucidate the population structure of *A. gambiae* are presented.

2. *Anopheles gambiae* species complex

Among the global vectors of human malaria arguably the most important species belong to the *Anopheles gambiae* complex, which include the most widespread and potent vectors of malaria in sub-Saharan Africa. The *Anopheles gambiae* species complex includes eight sibling species: *A. gambiae* s.s. Giles, *A. arabiensis* Patton, *A. bwambae* White, *A. melas* Theobald, *A. merus* Dönitz, *A. quadriannulatus* Theobald, *A. amharicus* Hunt, Coetzee and Fettene and *A. comorensis* Brunhes, le Goff and Geoffroy [1-4]. The status of these species was established via the demonstration of F_1 hybrid sterility among crosses between populations [4-8], morpho-

INTECH
open science | open minds

© 2013 Lanzaro and Lee; licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

logical features [9] and fixed differences in chromosomal inversions [5, 10]. Although the species cannot be reliably distinguished morphologically they do differ in terms of their ecology and geographic distributions (Figure 1). Two species, *A. merus* and *A. melas*, are associated with saltwater larval habitats and so are restricted in distribution to brackish water breeding sites along the east and west coasts respectively. A third saltwater species, *A. bwambae*, is only known to occur in association with hot springs in Semliki Forest National Park in eastern Uganda. The species, *A. quadriannulatus* and *A. amharicus* are primarily zoophilic and are not considered to be involved in the transmission of malaria. *A. quadriannulatus* occurs in southeastern Africa and *A. amharicus* in Ethiopia [2, 4]. A population on the island of Grande Comore in the Indian Ocean was described as a distinct species, *A. comorensis*, on the basis of morphological characters [9]. Little is known about the biology of *A. comorensis*. The two remaining freshwater species, *A. gambiae sensu stricto (*hereafter referred to as *A. gambiae)* and *A. arabiensis*, have the broadest geographic distribution and are the most important vectors of human malaria (Figure 1) [11, 12]. *A. gambiae* has been the most studied with respect to molecular and population genetics, and its whole genome sequence was published in 2002 [13].

Natural populations of *A. gambiae* have an extremely complex genetic structure that has been the subject of a great deal of research, a summary of which will be the focus of this chapter. Populations of *A. gambiae* are thought to be undergoing speciation and have been the focus of numerous studies aimed at evaluating speciation models [14-16]. Discrete subpopulations of *A. gambiae* have been defined in two ways: *chromosomal form* and *molecular form*. Recently the M molecular form of *A. gambiae* was elevated to species status and designated *Anopheles coluzzii* Coetzee *et al.* [4]. We retain the designation M and S forms to facilitate discussion of the recent literature.

3. Chromosomal forms of *Anopheles gambiae*

Chromosomal forms. The *A. gambiae* genome is organized on three chromosomes: two submetacentric autosomes and X/Y sex chromosomes, with males being the heterogametic sex. For descriptive purposes the autosomes are divided into two "arms" at the centromere. The longer arm is referred to as the right arm and the shorter the left arm. A high degree of chromosomal polymorphism, in the form of paracentric inversions, has been described in populations of *A. gambiae.* In a recent study Pombi *et al*. [18] describe 82 rare and 7 common inversions observed in natural populations. Inversions are not randomly distributed among chromosomes, but occur most often on the right arm of chromosome 2 (2R). Cytogenetic analysis is facilitated by the presence of giant polytene chromosomes in the cells of certain tissues. In early studies, the salivary glands of larvae were the source of material, but more recently ovarian nurse cells are used (the latter are easier to obtain and make better preparations for microscopic examination). Polytene chromosomes contain light and dark banding patterns that serve as critical landmarks for the determination of karyotypes (Figure 2). Protocols for the preparation of polytene chromosomes for karyotyping are available on-line at [19].

http://dx.doi.org/10.5772/56232

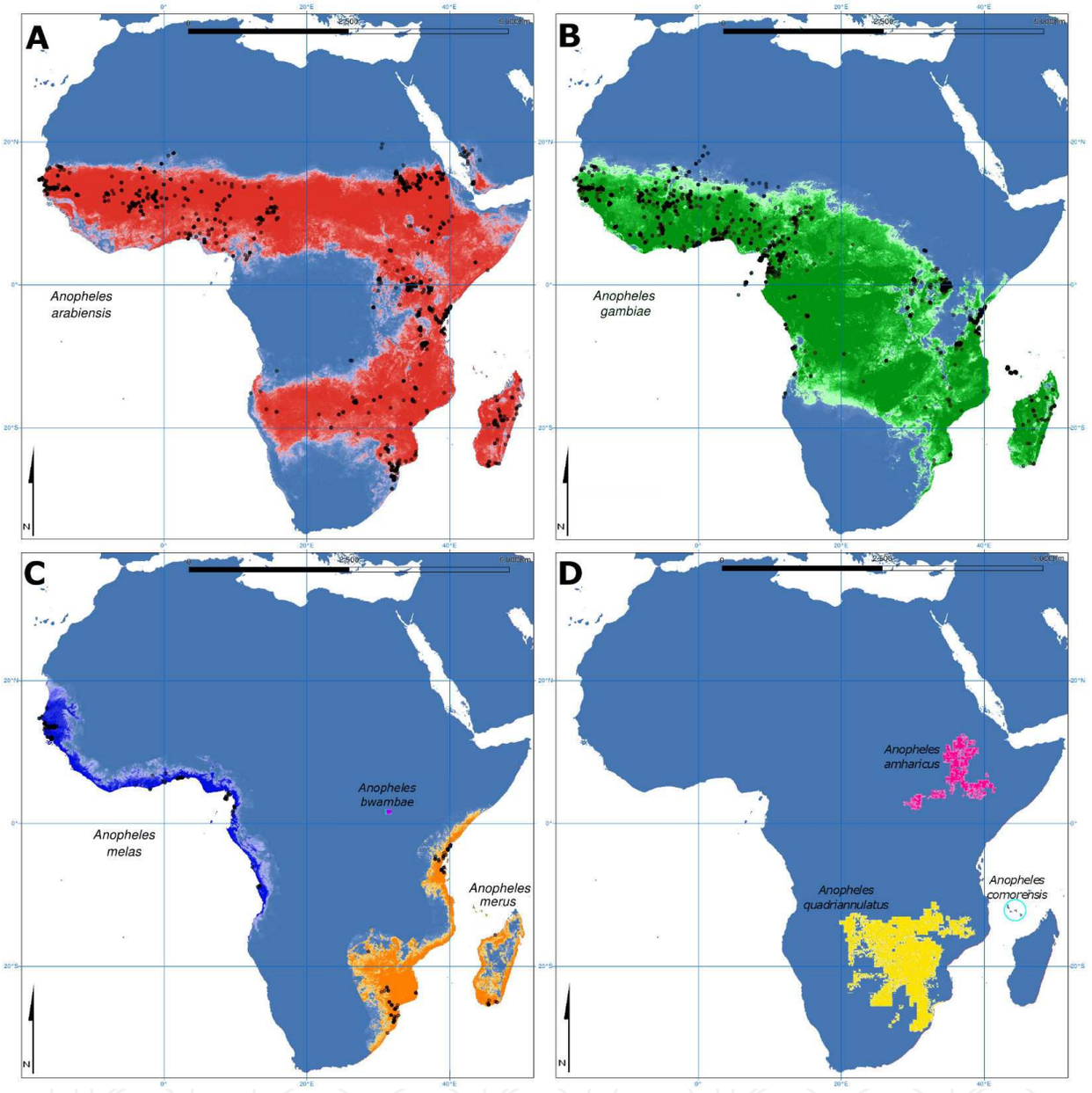


Figure 1. Geographic distribution of members of the *A. gambiae* complex. A: *A. arabiensis* (red); B: *A. gambiae s.s.* (green); C: *A. melas* (Blue), *A. merus* (orange), and *A. bwambae* (cyan); D: *A. quadriannulatus* (former species A) (yellow), *A. amharicus* (former *A. quadriannulatus* B) (magenta) and *A. comorensis* (cyan circle). Data and maps adapted from [17] and [14].

There is general agreement that inversions represent coadapted gene complexes that may enable individuals carrying them to occupy different ecological niches. The nonrandom distribution of inversion breakpoints along the chromosomes [18] and the distribution of inversion frequencies throughout the geographical ranges of the species strongly suggest that at least some of the inversions are maintained by selection that allows different species and, in the case of *A. gambiae,* populations, to survive and exploit a wide variety of habitats [21-23]. The best example is the strong association of inversions 2La and 2Rb with aridity, with the

frequency of these inversions being highest in drier areas and even increasing in frequency during the dry season at single sites that experience distinct wet and dry seasons [21, 23, 24]. Specific inversion configurations are associated with specific habitats, leading to the term "ecophenotype" frequently applied to describe individuals carrying certain combinations of inversions [25]. Chromosomal forms have been defined based on the configuration of five paracentric chromosome inversions on the right arm of chromosome 2 (2Rj, b, c, d and u) and one on the left arm of chromosome 2 (2La). Based on this, five chromosomal forms of *A. gambiae* have been described and named *Mopti*, *Bamako*, *Bissau*, *Forest* and *Savanna* according to the geographic regions from which they were first collected and indicating an association of each with a particular type of habitat, as illustrated in Figure 3 [10]. Chromosomal forms are defined as follows: [1] the *Forest* form characterized by the typical non-inverted arrangement 2R+/+, 2L+/+, or by a single inversion polymorphism due to inversion 2Rb, 2Rd or 2La; [2] *Bissau* characterized by high frequencies of the 2Rd inversion and standard 2L+ arrangement; [3] *Savanna* exhibiting high frequencies of 2Rb and 2La inversions as well as polymorphism involving the 2Rcu arrangements and polymorphism in the j, d and the rare k inversion; [4] *Bamako* characterized by the fixed 2Rjcu arrangement and polymorphism in the 2Rb inversion; [5] *Mopti* showing high frequencies of 2Rbc, 2Ru and nearly fixed for 2La (Figure 2). The *Savanna* form has the broadest distribution occurring throughout sub-Saharan Africa, the *Mopti* form predominates in drier habitats in West Africa, the *Forest* form occurs in wetter habitats in Africa, the *Bamako* form occurs in habitats along the Niger River in West Africa and the *Bissau* form is restricted to West Africa (Figure 3) [26, 27].

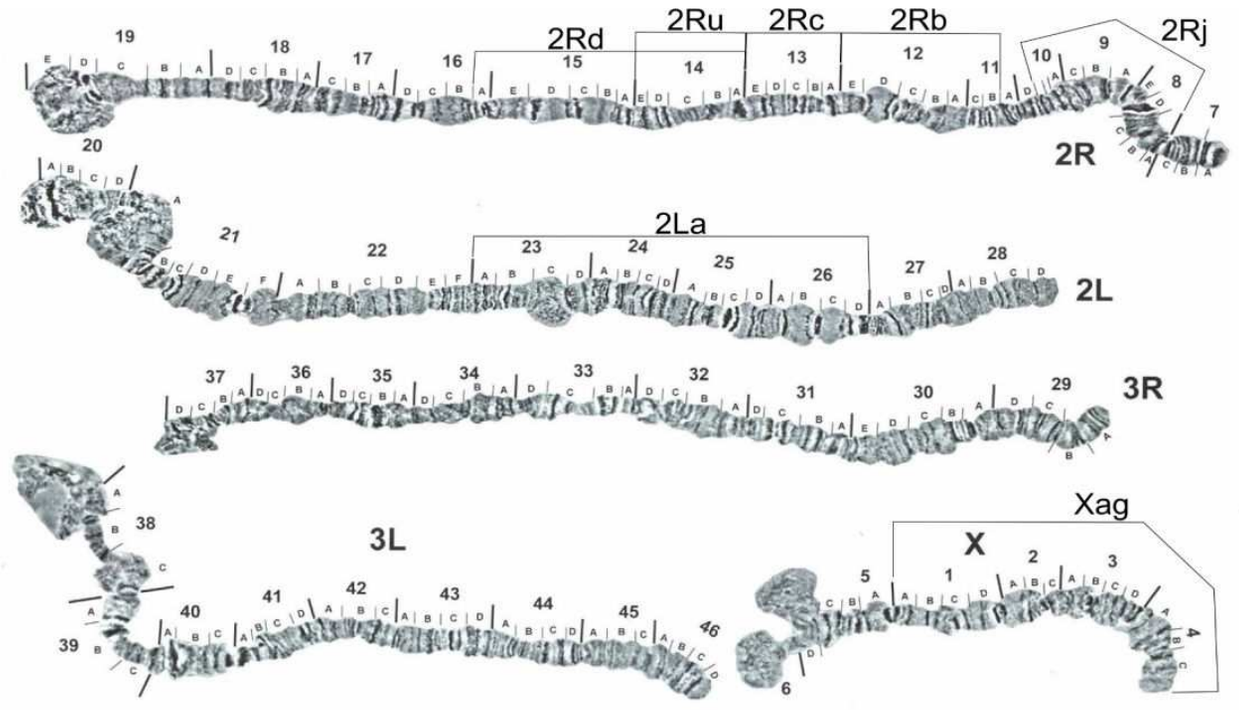


Figure 2. Photomap of polytene chromosomes of *A. gambiae* Forest-M form (collected from Tiko, Cameroon) depicting band positions. Six major inversions on the chromosome 2 used for identifying chromosomal forms are marked.

http://dx.doi.org/10.5772/56232

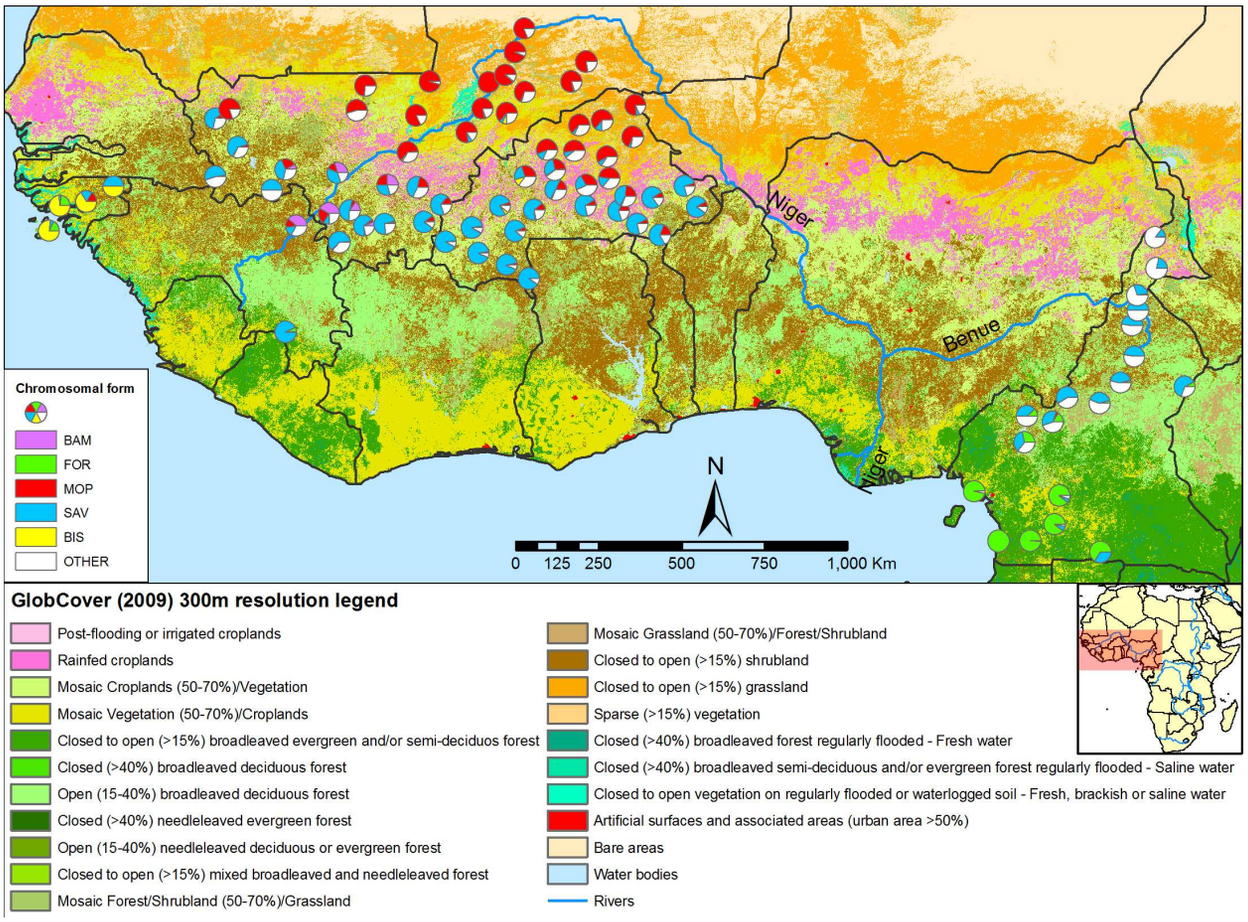


Figure 3. Distribution of chromosomal forms in West and Central Africa. Data from *PopI* [27]. BAM stands for *Bamako* chromosomal form, FOR for *Forest*, MOP for *Mopti*, SAV for *Savanna* and BIS for *Bissau*. OTHER refers to samples with karyotypes that do not fit any described chromosomal form designation.

The Xag inversions is fixed and used as a diagnostic marker to distinguish *A. gambiae* from other species in the complex. Chromosome photomap adapted from: [20]

It has furthermore been suggested that the chromosomal forms are to some extent reproductively isolated and represent distinct species or incipient species that have evolved or are evolving via a process described as "ecotypic speciation" [15, 25]. Studies of karyotype frequencies at sites where the *Bamako*, *Mopti* and *Savanna* forms occur in sympatry have revealed significant departures from the Hardy-Weinberg equilibrium (H-W) [10, 28-30]. Specifically, heterokaryotypes representing hybrids between the *Savanna* form and the other two were under-represented and *Bamako*/*Mopti* hybrids were never encountered. This observation led to the suggestion that there is partial reproductive isolation between *Savanna* and the other forms, nearly complete isolation between the *Bamako* and *Mopti* forms and that these forms represent incipient species. However, hybridization experiments involving crosses between the *Bamako* and *Mopti* forms resulted in viable offspring, demonstrating a lack of post-mating reproductive barriers between them [29, 31]. An estimate of genetic distance (based on allozyme frequencies) [32] between the *Bamako* and *Mopti* forms was reported as 0.015 [33], a value not higher than that typically found between local populations of a single

mosquito species. We found that genotypic frequencies in a population composed of three chromosomal forms in Mali did not depart from Hardy-Weinberg expectations, suggesting that this population represents a single gene pool (Lanzaro, unpublished).

It should be emphasized that although these studies do not support reproductive isolation among chromosomal forms, they do not disprove it. Pre-mating isolating mechanisms may act as a barrier between subpopulations, even if post-mating mechanisms have not evolved, and isolation may be recent, so that not enough time has passed for the accumulation of substantial allozyme divergence between the forms. Lanzaro *et al.* [34] conducted a study based on 21 microsatellite loci distributed over the genome, examining genetic differentiation between the *Bamako* and *Mopti* forms in the villages of Banambani and Selinkenyi, Mali. This study revealed strong genetic differentiation between *A. gambiae* and *A. arabiensis*, used here as an outgroup. Within *A. gambiae*, different patterns of genetic differentiation, depending on the genomic location of the microsatellite loci, were observed. No genetic differentiation was found on the 3rd and X-chromosome whereas strong linkage disequilibrium and low levels of genetic differentiation were found for loci located on the 2nd chromosome in association with the inversions that occur there [34]. Similar results were obtained in a study also using microsatellites distributed on all three chromosomes for samples collected in the villages of Selinkenyi, Soulouba, and Kokouna, Mali [35].

Gene flow, like other forces, may be higher in some parts of the genome and lower in others. For example, favorable genes can still be exchanged successfully even when barriers to gene flow are strong. Such genes could be at loci that confer local adaptations and at any linked loci. The significance of this is that gene flow, even if estimated accurately, may still fail to account for variation among different parts of the genome. This effect may be particularly strong for genes contained within inversions, both because of potentially strong selection and because of linkage imposed by the reduced recombination associated with inversions. This effect was explored by Tripet et al. [36] in a study in which they examined divergence for microsatellite loci contained within the *j* and *b* inversions compared with loci outside of inversions. Indeed they did find elevated divergence estimated from loci contained within the inversions relative to those outside. This pattern of divergence, with a strongly non-random distribution over the genome, was later described as a 'mosaic genome architecture' in a paper by Wang-Sattler et al. [37]. As we shall see, this concept was later refined based on high resolution genome-wide analysis, ultimately leading to the recognition of 'islands of speciation' in the *A. gambiae* genome.

Using the chromosomal form concept to define genetically discrete populations is problematic because there is substantial overlap in inversions that define them, probably due to some level of contemporary gene flow. This creates ambiguities in assigning individuals to form, diminishing the utility of the chromosomal form concept for defining reproductive boundaries among populations. For example, in a recent survey of populations in Mali, we found that 26% of 2,459 individuals could not be assigned to a chromosomal form and in Cameroon 39% of 632 individuals could likewise not be assigned (Figure 3, data available at *PopI* [27]).

The role of chromosome inversions in A. gambiae evolution: Ecotypic Speciation. The chromosomal or ecotypic model of speciation was first described for anopheline mosquitoes by Coluzzi [38]

http://dx.doi.org/10.5772/56232

and is the prevailing model applied to the chromosomal forms of *A. gambiae* [14, 15]. This model is founded on the observation that certain paracentric inversions that are polymorphic in *A. gambiae* are non-randomly distributed in nature. These are thought to contain multi-locus genotypes that are adaptive to specific aquatic habitats occupied by the immature stages of the mosquito. Under this model, populations carrying alternate gene arrangements would inhabit different, spatially isolated habitats. Genetic divergence, enhanced by reduced recombination associated with the inversions, would then evolve. Ultimately divergence would include genes resulting in reproductive isolation (reduced fitness in hybrids or behavioral differences preventing between form mating), explaining the observed deficiency of inversion heterozygotes. This model was initially adopted to describe the evolution of chromosomal forms of *A. gambiae* [15, 21, 28], but now has become the model for explaining the evolution of the molecular forms as described below [16, 39-42].

The most thorough evaluation of the ecotypic speciation model has been its application to the *Bamako* and *Savanna* forms in Mali [15]. Central to this evaluation is the observation of niche partitioning with respect to larval habitat. This observation was based on a PCR identification method developed for detecting the 2R*j* inversion [43] among larval samples collected in rock pools vs. more typical larval sites (puddles/ponds) in the village of Banambani, Mali. We evaluated this PCR method on a set of 85 field-collected adults previously scored for the 2R*j* genotype cytogenetically. In total, we selected 25 2R*j* homozygotes (*j*/*j*), 40 2R*j* heterozygotes ($+_j/j$) and 20 2R*j* standard ($+_j/+_j$) from the villages of Banambani, Selinkenyi, Tinko and Seroume, Mali. The 2R*j* PCR was accurate in calling 2R*j* homozygotes (*j*/*j*) (100%) in all villages regardless of the presence of the 2R*c* and *u* inversions (Table 1). However, the PCR was much less accurate for the standard arrangement for 2R*j* ($+_j/+_j$), resulting in consistent false identification as 2R*j* heterozygotes ($+_j/j$) in 11 cases and 2Rj (*j*/*j*) homozygotes in 5 cases. Moreover, all true heterozytoes ($+_j/j$) were misidentified as either *j*/*j* (N=13) or $+_j/+_j$ (N=7). The low accuracy rate (=48.2%) of the 2R*j* diagnostic PCR casts doubt on this sole example of niche partitioning (rock pool vs. other) in larval habitat distinguishing the two forms.

The 2R*j* inversion polymorphism in Mali shows two mating patterns in different parts of the species range in this country. At sites along the Senegal River (e.g. villages of Sebetou, Seroume, Bantinngoungou, and Tinko), 2R*j* inversion heterozygotes are commonly found and 2R*j* karyotypes are in Hardy-Weinberg expectation (HWE). On the other hand, at sites along the Niger River and its tributaries (e.g. villages of Banambani, Doneguebougou, Senou, Kela, Selinkenyi, Soulouba, Yorobougoula, Kokouna), a severe deficiency of 2R*j* heterozygotes are observed and 2R*j* genotypes are not in HWE (Figure 4).

In the literature the *Bamako* form includes three genotypes, *jcu*/*jcu*, *jcu*/*jbcu*, and *jbcu*/*jbcu*, all homozygous for *j* [21]. Other individuals carrying 2R*j* inversion but not *c* and *u* inversions such as *jbd*/*jbd*, and *jb*/*b*, commonly found along the Senegal River, cannot be classified under the current definitions for chromosomal forms. 94% of the 2R*j* homozygotes along the Niger River are *Bamako* forms, while no *Bamako* forms are found along the Senegal River.

Overall these results weaken the argument that paracentric inversions play a role in the evolution of reproductive isolation via divergent selection (ecotypic speciation), both because

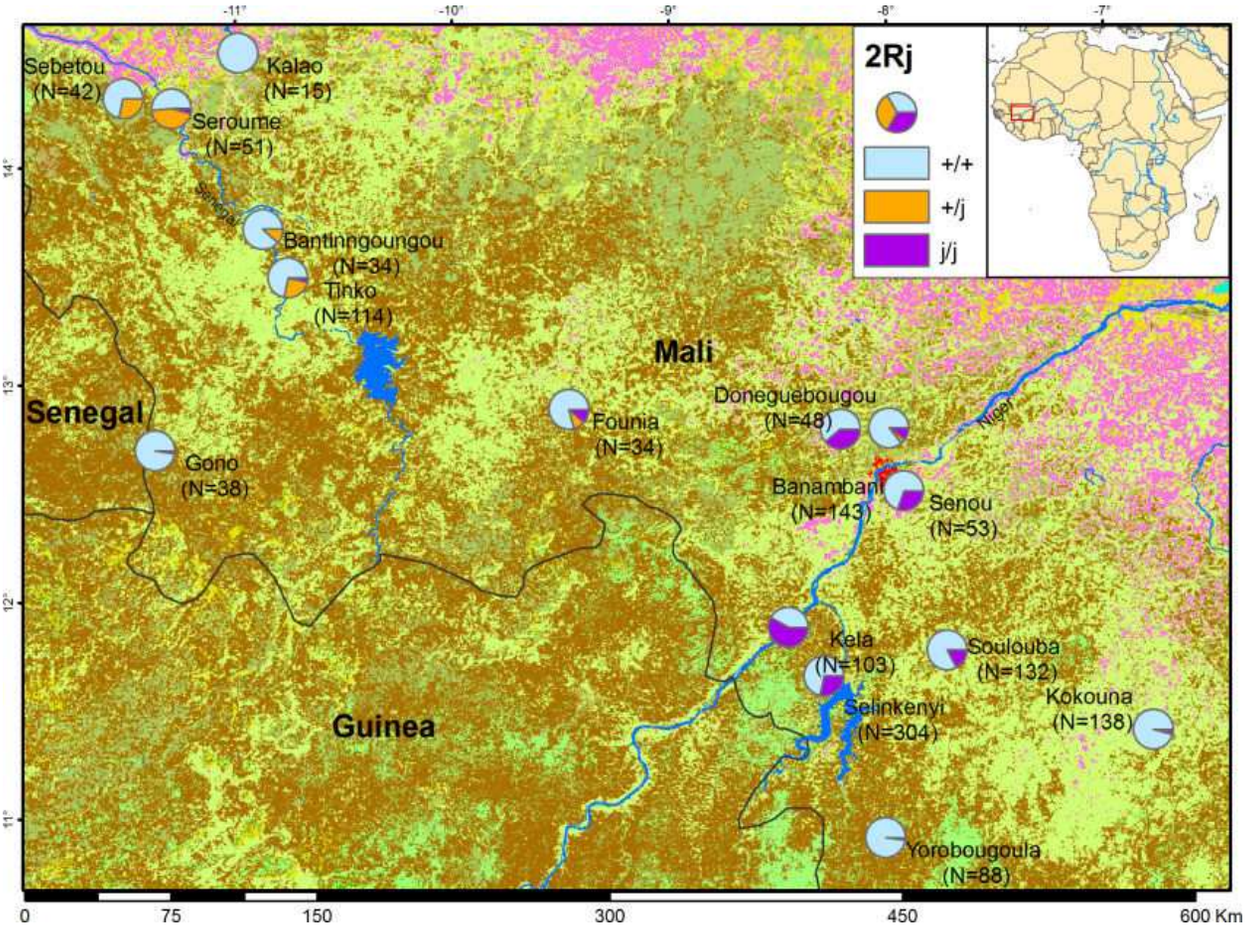


Figure 4. 2R*j* inversion distribution in Mali. For legend of the GlobCover 2009 land cover type used as background, see Figure 2.

		Run 1			Run 2		
2Rj	**N**	**match**	**mismatch**	**NA***	**match**	**mismatch**	**NA***
j/j	25	24	0	1	24	0	1
$+_j$/j	20	0	**20**	0	0	**19**	1
$+_j$/$+_j$	40	17	**14**	9	16	**16**	8

Table 1. Evaluation of 2R*j* genotyping via a PCR identification method. Samples were karyotyped microscopically prior to being assayed using the PCR protocol of Coulibaly *et al*. [43] * NA stands for 'no amplification'

they cast doubt on the association of inversions with distinct larval habitats and on evidence for reproductive isolation between individuals that differ with respect to the inversions they carry (e.g. a lack of *j* inversion heterozygotes). Genome-wide comparisons of individuals with and without inversions have been conducted and these cast doubt on the role of inversions as forming "coadapted gene complexes". These results are described in detail below.

http://dx.doi.org/10.5772/56232

The role of chromosome inversions in A. gambiae evolution: Comparative Genomics. Central to the "ecotypic speciation" model as applied to *A. gambiae* is the notion that inversions contain multilocus genotypes that are adaptive to different environments. These "coadapted gene complexes" arise and are maintained as the consequence of reduced recombination within and around the inversion. Ultimately these become, either directly or indirectly, associated with reproductive isolation. One expectation arising from this phenomenon, assuming that reproductive isolation is incomplete or has evolved recently, is higher levels of genetic divergence in regions of the genome contained within the inversion relative to elsewhere in the genome. Indeed, in a genome-wide scan comparing individuals with and without the 2L*a* inversion, significantly higher divergence was observed in a 3 Mb region of the genome within and proximal to the inversion [44]. However, in a subsequent study that included a comparison of inverted and uninverted genomes for the four common 2R inversions (*j*, *b*, *c* and *u*), a region of the genome spanning ~26 Mb, divergence was limited to just one small region (~100 kb) in the 2R*u* inversion [45]. In both studies the Affymetrix *Plasmodium*/*Anopheles* Genome Microarray (*P*/*A* array), which contains 142,065 25bp probes, representing roughly 13,000 predicted genes, was used. Lack of divergence associated with the inversions hypothesized to be driving the "ecotypic speciation" process was unexpected. Several explanations were provided including that divergence between the inversion arrangements escaped detection due to shared ancestral polymorphism, extensive recombination within the inversions (gene flux) and limits to the resolution of the microarray they used [45].

In a more recent study [46] the genomes of individuals homozygous for the *jbcu* arrangement (*Bamako* form) were compared with individuals homozygous for the standard arrangement, $+_j+_b+_c+_u$ (*Savanna* form). In this case all individuals were of the S molecular form (unlike the comparisons made in the White et al. [45] study, which were a mixture of M and S form individuals). In addition, Lee et al. [46] utilized an *A. gambiae* whole genome tiling microarray (WGTM) which provides a far higher resolution of the genome than the *P*/*A* array (probe density = 1 probe per 100,000bp for the *P*/*A* array; 1 probe per 17bp for the WGTM). As in the White et al. [45] study, this new study revealed very little divergence associated with the chromosome 2R inversions. However, a 3Mb region of the genome on the X chromosome, proximal to the centromere was observed. This is the same region of the genome that contains the sequence divergence used to define the M and S molecular forms (discussed in detail in the following sections). X chromosome divergence is associated with reproductive isolation observed between both the M and S molecular forms and between the *Bamako* and *Savanna* chromosomal forms. These results suggest that the 2R inversions may not be involved in either the evolution or maintenance of reproductive isolation among *A. gambiae* populations.

4. Molecular forms of *Anopheles gambiae*

Defining Molecular Forms. An attempt to develop a molecular diagnostic for the chromosomal forms of *A. gambiae* identified 10 nucleotide residues that differ between the *Mopti* and the *Savanna* or *Bamako* chromosomal forms in a 2.3 kb fragment at the 5′ end of the rDNA IGS region located on the X chromosome [47]. These findings led to the development of a PCR-

based diagnostic to differentiate *Mopti* chromosomal forms from *Bamako* and *Savanna* forms based on a single base pair substitution at the 540th nucleotide position in a 28S rDNA amplimer sequence. *Mopti* form individuals carry a C/C genotype and both *Bamako* and *Savanna* individuals a T/T genotype (Genbank accession number AF470112-6) [48]. Individuals carrying C/C are referred to as M molecular form and those carrying the T/T genotype are known as S molecular form. There is good correspondence between the M molecular form and the *Mopti* chromosomal form in Burkina Faso and Mali, however, the *Bamako* and *Savanna* chromosomal forms cannot be distinguished (both are of the S molecular form). The association of M and S molecular forms and chromosomal forms breaks down at other locations in West Africa. For example, in western Senegal and Gambia the association between the *Savanna* chromosomal form and S molecular form does not hold [49] and the *Forest* form contains both M and S individuals. The M and S molecular forms, therefore, largely fail as a diagnostic for chromosomal form. However, the significance of the M and S forms of *A. gambiae* goes well beyond their utility as proxies for identifying chromosomal forms. The molecular form concept has now largely replaced chromosomal form for defining discrete sub-populations of *A. gambiae*, that are to some extent reproductively isolated.

M and S forms occur in sympatry at many sites in West and Central Africa, and typically there is a high degree of reproductive isolation between the two forms. M/S hybrids (C/T genotype) produced in the laboratory did yield clearly distinguishable hybrid patterns in females. Surprisingly, however, field collected individuals carrying "hybrid" karyotypes (putative hybrids between different chromosomal forms) did not produce results consistent with their being hybrid, but rather produced either M or S patterns [48]. This observation supports the notion that certain karyotypes, thought to be fixed in one chromosomal form or another, are in fact shared, occurring commonly in one form and rarely in another, due to ancestral polymorphism and/or ongoing gene flow [40, 50]. This diagnostic now forms the basis of recognizing two distinct subpopulations of *A. gambiae*, known as *molecular forms* (M and S).

Alternate methods for distinguishing M and S forms. The original PCR-based diagnostic used to distinguish the M and S forms [48] was further developed into a method using a restriction digestion of PCR amplimers that allowed distinguishing *A. gambiae* from one of its sibling species *A. arabiensis* while simultaneously distinguishing M from S [51]. This was useful in the field since *A. arabiensis* and both the M and S forms are morphologically indistinguishable and commonly occur in sympatry at study sites throughout West and Central Africa. In 2008, a new method for distinguishing the M and S forms was discovered which takes advantage of polymorphism in insertion sites for a group of retrotransposons known as short interspersed elements (SINEs). One of the SINE insertion sites, located on the X chromosome and referred to as SINE X6.1, was found to be fixed in the M form and absent in the S form. In subsequent studies, in which multiple M/S diagnostic methods were employed, some discrepancies in results were observed [52]. These were most common in populations where M/S hybridization is common, for example in Guinea-Bissau.

Relationships between the M and S forms. Understanding the relationship between the two molecular forms has been the focus of an intense and ongoing research effort. The S form has the broadest distribution occurring throughout sub-Saharan Africa, whereas the M form

http://dx.doi.org/10.5772/56232

occurs throughout West and parts of Central Africa. With the exception of a single site in northern Zimbabwe [53], M is absent from eastern Africa (Figure 5) [49].

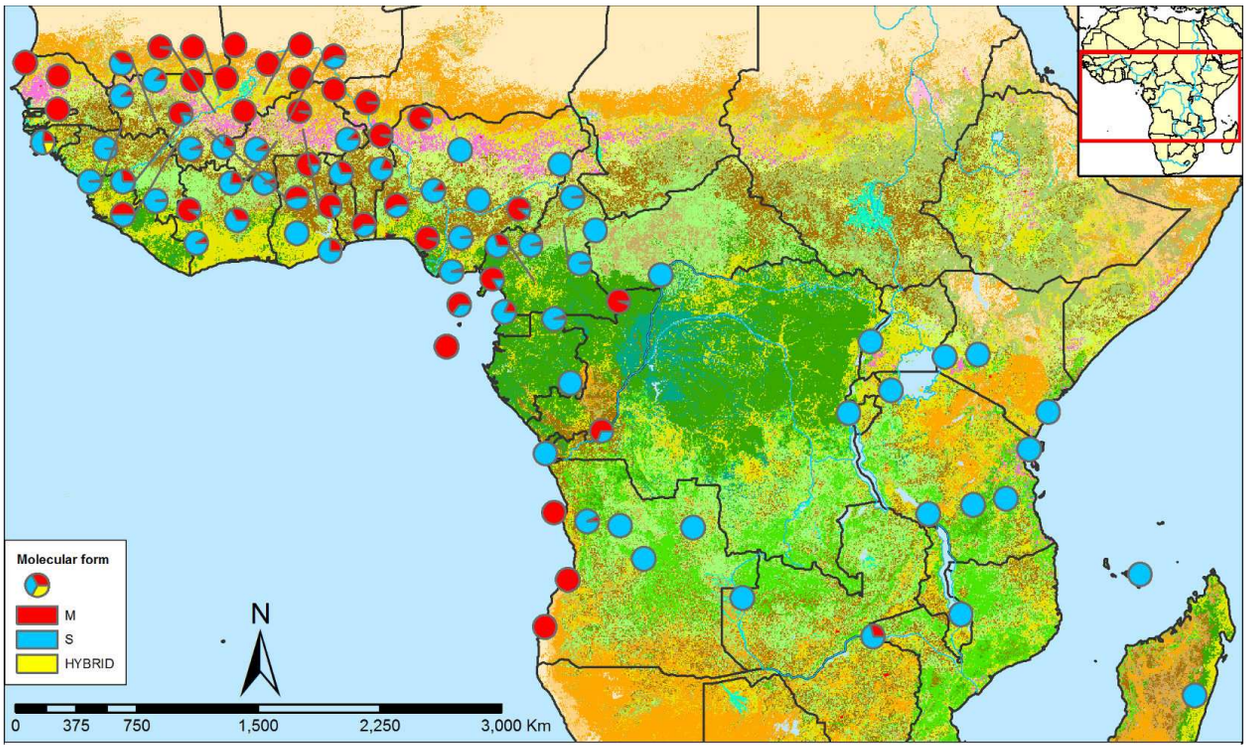


Figure 5. Distribution of molecular forms in Sub-Saharan Africa. For legend of the GlobCover 2009 land cover type used as background, see Figure 2. Data from [27, 40, 49, 50, 54]

Although the M and S forms are largely reproductively isolated in most places where they occur together, this is not true everywhere. Hybridization between forms occurs rarely (~1%) in Mali [55] and reproductive isolation between M and S appears to be complete in Cameroon [56]. In The Gambia, M/S hybrids were identified from a number of sites at frequencies as high as 16.7% of the *A. gambiae* individuals sampled [57] and in Guinea-Bissau hybrids were recovered in over 20% of the individuals assayed [58, 59]. A cryptic subgroup of *A. gambiae* known as the "Goundry" population collected in Burkina Faso was recently found to be composed of 36% M/S hybrids [60]. The Goundry population discovered in the Sudan Savanna zone of Burkina Faso in larval collections but absent in indoor adult collection of the same locality, suggesting that adult stage of Goundry populations mostly rest outdoors [60, 61]. These results suggest that linkage between the M and S alleles and those genes that directly affect reproductive isolation has broken down in a much broader geographic area than previously suggested. Therefore, the notion of an M form and an S form that are largely reproductively isolated (incipient species) and that hybridization only occurs in the "Far-West" region of Africa [62] is an oversimplification.

In the laboratory, chromosomal and molecular forms, including the *Bamako* and *Savanna* forms, appear to display no post-zygotic isolation [31, 63, 64]. Analysis of sperm recovered from inseminated females [55] and the composition of mating swarms [65] support the existence of

strong, but not complete, pre-mating reproductive isolation between the M and S molecular forms in nature.

The two molecular forms display phenotypic divergence in different locations within their geographic range [66]. Most notable among these phenotypic differences include differential insecticide resistance [67], desiccation resistance [68], larval habitat segregation [69], and wing morphological differentiation [70]. It has been proposed that the mechanism responsible for promoting divergence is pre-zygotic [63] and associated with mate selection either during swarm formation [71, 72] or within a swarm [65]. Diabate et al. found evidence of clustering of swarms composed of individuals of a single molecular form within the village of Donégué-bougou, Mali [71]. Mixed swarms of M and S forms were found elsewhere (Burkina Faso) but the occurrence of mixed swarms was lower than the frequency expected by chance. Manoukis *et al.* analyzed the shape of male swarms and suggested that a difference in swarm organization between M and S forms may enhance the behavioral isolation of the two forms [72].

5. Evolution of the M and S forms

Comparative genomics. Early studies aimed at describing patterns of genetic divergence among chromosomal forms revealed what was termed a "mosaic genome architecture", with divergence distributed non-randomly over the genome (as described above, [29]). Comparisons of the M and S forms revealed a similar pattern. Initial work examined the distribution of microsatellite DNA polymorphism showing exceptionally high divergence in a region of the genome proximal to the centromere on the X chromosome, near the rDNA locus used to define the two forms [35, 73]. High levels of M/S form divergence on the X chromosome was substantiated through detailed examination of the centromeric region using DNA sequencing [74, 75].

The first high density genome-wide comparison of M and S was conducted by Turner et al. [76] using samples collected in Cameroon. They utilized an Affymetrix *Plasmodium/Anopheles* Genome Microarray which contains 142,065 25bp probes representing roughly 13,000 predicted genes. Divergence between the M and S genomes was very low and restricted to three discrete regions, one on the X chromosome (corresponding with the location identified in the microsatellite studies) and two on chromosome 2, one on 2L and one very small (37kb) region on 2R. In total, these diverged regions cover less than 2.8Mb, roughly 1% of the genome. In a subsequent study, utilizing the same microarray, but with samples collected in Mali, the small 2R region of divergence was not observed, and so this small region was considered not to contribute to reproductive isolation between the two forms [77]. Later a third diverged region was observed on the left arm of chromosome 3L and this region, like the X and chromosome 2L regions, was proximal to the centromere [16]. Taken together these studies revealed that the M and S genomes are diverged over only about 3% of their genomes and that this divergence is organized into 3 small regions located near the centromere on the X, 2L and 3L chromosomes, with the remainder of their genomes essentially undifferentiated. These regions

http://dx.doi.org/10.5772/56232

of divergence have been considered to represent *islands of speciation* because it is thought that they contain genes that are directly involved in reproductive isolation.

Islands of speciation model. The widely held interpretation of this work is that *A. gambiae* forms represent incipient species, but with enough gene flow to prevent their genomes from diverging in all but a few, relatively small regions [34, 35, 37, 76, 77]. This interpretation is consistent with recent genic models of speciation that predict the existence of small regions of divergence between incipient species in the presence of some degree of gene flow (Figure 6) [78, 79]. The observation that putative "*islands of speciation*" in *A. gambiae* are located proximal to centromeres, where levels of recombination are known to be low, is likewise consistent with models that consider speciation to be driven by genes located in regions of the genome with reduced crossing-over [80, 81].

Incidental islands model. White et al. [16] developed PCR-RFLP assays to detect SNPs that occurred in each of the three islands of speciation and that were diagnostic for the M and S forms. They genotyped a total of 517 individuals including both M and S forms from Mali, Burkina Faso, Cameroon and Kenya. They found complete association among the three unlinked islands in 512 of the 517 individuals genotyped (275 M form and 237 S). Of the five exceptional genotypes, three were heterozygous at all three loci, suggesting these represented F_1 hybrids. To account for the nearly complete linkage between the three diverged islands they suggest that gene flow between M and S must be nearly zero. The presence of F_1 hybrids suggests that they have such low fitness that they contribute little to gene flow between the forms. As mentioned above F_1 hybrids generated in the laboratory show no evidence of intrinsically low fitness, so it is assumed that these are maladapted to conditions in nature. Additional support for very low levels of between form gene flow come from comparisons of M and S based on high-density, genome-wide SNP genotyping [41] and whole genome sequences [42] which revealed widespread divergence between the M and S genomes. Collectively these studies propose an alternative model referred to as the "incidental islands" model [82, 83], which states that reproductive isolation between M and S is complete and that the observed islands of divergence may be incidental, meaning that the divergence observed in areas proximal to centromeres do not necessarily represent the location of genes underlying reproductive isolation but the divergence is due to segregating ancestral variation and not due to contemporary gene flow.

In summary, two opposing models exist that describe the relationship between the M and S forms. The "genomic islands of speciation" model suggests that divergence between the M and S genomes is restricted to small regions (~3% of the genome) that may contain the genes responsible for reproductive isolation between forms and that ongoing gene flow is responsible for very low levels of divergence over the remaining 97% of the genome. The second model, the "incidental islands of divergence" model, suggests that divergence between the two forms is far more extensive and widely distributed over the genome, that gene flow between the two forms is nearly zero and that the M and S forms therefore represent distinct species (Figure 6D).

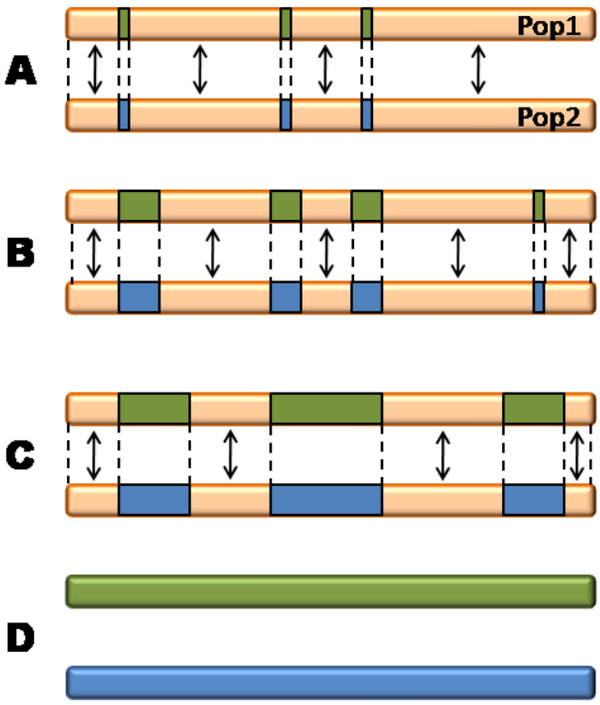


Figure 6. A: Stage 1 - Population/races with differential adaptation; reproductive isolation (RI) not apparent. Green box represents diverged loci specific to Population 1 (Pop1) and Blue represents diverged loci specific to Pop2. Arrows indicate regions of gene flow. B: Stage 2 - Transition between races and species with some degree of RI; population may fuse or diverge. C: Stage 3 - Divergent populations beyond the point of fusion but still share a portion of their genome via gene flow; good species. D: Stage 4 - Species with complete RI. Adapted from [79].

6. Further sub-divisions within molecular forms

Although most discussions consider M and S as the major and biologically relevant subdivisions of *A. gambiae* there is evidence that the two can be further subdivided into population groups that are significantly diverged.

Subdivision within the S form. In a continent-wide survey Lehmann et al. [73] found that S form populations fall into two well defined clades, based on analysis of microsatellite DNA. They refer to these clades as the Northwest (Nigeria, Gabon, Democratic Republic of Congo, NW Kenya) and Southeast (SW Kenya, Tanzania, Malawi) divisions. Wang-Sattler et al. [37] also conducted an analysis based on microsatellite DNA and likewise report that the S form in eastern Africa (Kenya) are distinct from S form populations in the west (Mali). In addition to the East vs. West division between allopatric S form populations is the division of sympatric S form populations in Mali. These are described in detail above (Section 2). In brief, the S form in Mali is divided into the *Bamako* and *Savanna* chromosomal forms which display strong asssortative mating where they occur in sympatry at sites along the Niger River ([21], also see Figure 4). These two populations can be distinguished by the *j* inversion, which is fixed in the *Bamako* form and absent in the *Savanna*. Interestingly, although the two share the X-linked

http://dx.doi.org/10.5772/56232

allele that defines them as S molecular form, a detailed analysis revealed that they are strongly diverged at a 3Mb region of the X chromosome, proximal to the centromere [46].

Subdivision within the M form. A comparison of the M form in Mali and the M form in Cameroon has revealed that the two are very different genetically, in fact, divergence between these two is higher than the level of divergence between the M and S forms [23]. This observation has led to a recognition of two, distinct M form groups, the Mopti-M form, which is polymorphic with respect to the 2R *b*, *c*, and u chromosome inversions and the Forest-M form which lacks inversions on chromosome 2L and 2R [23, 84]. In addition to genetic divergence the Forest-M and Mopti-M forms differ in their ecology. The Mopti-M in Mali is most common in the dry northern part of the country whereas Forest-M is absent in the dry northern part of Cameroon and is restricted to the wet southern part of the country [23]. This observation lends support the notion that chromosome inversions are involved in adaptation to arid environments.

The Goundry form. Genetic analysis of *A. gambiae* larvae from roadside pools in Burkina Faso and adults collected from inside nearby houses revealed the occurrence of a genetically distinct population present in the larval sample, but absent from adult collections [60]. The larval population differed from the adult population with respect to the distribution of microsatellite alleles (F_{ST}=0.15), the presence of M/S hybrids (35% in the larval population, <1% in adults) and in the frequency of the 2L*a* inversion (2L*a* = 58% in larval population, 96% in adults). This distinct larval population is called the Goundry form, after one of the village collection sites. Based on these results it is supposed that the Goundry form is a unique form in which the adults rest nearly exclusively outdoors (exophilic) and which, although they carry the X-linked genetic markers that distinguish the M and S forms, the assortative mating associated with these markers is absent. Adults of the Goundry form have never been collected. Adults reared from larvae of the Goundry form were found to have increased susceptibility to infection with *P. falciparum* in laboratory experiments. [60]

7. Future directions

Reconciliation of the opposing speciation models and clarification of new "forms" await the resolution of a number of outstanding questions concerning interactions between the M and S forms. It is clear that the determination of the frequency of hybrid individuals requires that individuals be identified using multi-locus genotypes at unlinked loci, such as those employed by White et al. [16], as opposed to the widely used single locus X-linked markers. This would allow not only the recognition of F_1 hybrids but backcross individuals as well. Determination of the frequencies of both F_1 and backcross genotypes would provide information on the level of introgression. Moreover, multi-locus approach will allow identification of hybrid males. The application of this method to populations throughout the sympatric range of M and S would allow a description of spatial heterogeneity in levels of introgression that could be related to key environmental parameters that include mating cues that sustain assortative mating within forms as well as conditions that favor the survival of hybrid genotypes.

Author details

Gregory C. Lanzaro* and Yoosook Lee

*Address all correspondence to: gclanzaro@ucdavis.edu

Department of Pathology, Microbiology and Immunology School of Veterinary Medicine University of California Davis, CA, USA

References

[1] White GB. *Anopheles bwambae* sp.n, a Malaria Vector in the Semliki Valley, Uganda, and Its Relationships with Other Sibling Species of the *An. gambiae* Complex (Diptera, Culicidae). Syst Entomol. 1985;10(4):501-22.

[2] Hunt RH, Coetzee M, Fettene M. The *Anopheles gambiae* complex: a new species from Ethiopia. Trans R Soc Trop Med Hyg. 1998 Mar-Apr;92(2):231-5. PubMed PMID: 9764342. Epub 1998/10/09. eng.

[3] Harbach R. Mosquito Taxonomic Inventory http://mosquito-taxonomic-inventory.info/2012 [cited 2013]. Available from: http://mosquito-taxonomic-inventory.info/sites/mosquito-taxonomic-inventory.info/files/Anopheles%20classification_25.pdf.

[4] Coetzee M, Hunt RH, Wilkerson R, della Torre A, Coulibaly MB, Besansky N. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. Zootaxa. 2013;3619(2):246-74.

[5] Davidson G, Hunt RH. The crossing and chromosome characteristics of a new, sixth species in the *Anopheles gambiae* complex. Parassitologia. 1973 Apr-Aug;15(1):121-8. PubMed PMID: 4788353. Epub 1973/04/01. eng.

[6] Davidson G. Insecticide resistance in *Anopheles gambiae* Giles: a case of simple mendelian inheritance. Nature. 1956 Oct 20;178(4538):863-4. PubMed PMID: 13369563. Epub 1956/10/20. eng.

[7] Davidson G. The Five Mating-Types in the *Anopheles gambiae* Complex. Riv Malariol. 1964 Dec;43:167-83. PubMed PMID: 14318975. Epub 1964/12/01. eng.

[8] Davidson G. *Anopheles gambiae*, a Complex of Species. Bull World Health Organ. 1964;31:625-34. PubMed PMID: 14278001. Pubmed Central PMCID: PMC2555133. Epub 1964/01/01. eng.

[9] Brunhes J, LeGoff G, Geoffroy B. Afrotropical *Anopheles*.1. Description of new species and modifications of taxonomical status (Diptera:Culicidae). Ann Soc Entomol Fr. 1997 Jun;33(2):173-83. PubMed PMID: WOS:A1997XM01900005. French.

[10] Coluzzi M, Sabatini A, Petrarca V, Di Deco MA. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. Trans R Soc Trop Med Hyg. 1979;73(5):483-97. PubMed PMID: 394408. Epub 1979/01/01. eng.

[11] Gillies M, De Meillon B. The Anophelinae of Africa south of Sahara. Johannesburg: The South African Institute for Medical Research, 1968 Contract No.: 54.

[12] Coetzee M, Craig M, le Sueur D. Distribution of African malaria mosquitoes belonging to the *Anopheles gambiae* complex. Parasitol Today. 2000 Feb;16(2):74-7. PubMed PMID: 10652493. Epub 2000/02/01. eng.

[13] Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, et al. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science. 2002 Oct 4;298(5591):129-49. PubMed PMID: 12364791. Epub 2002/10/05. eng.

[14] Ayala FJ, Coluzzi M. Chromosome speciation: humans, Drosophila, and mosquitoes. Proc Natl Acad Sci U S A. 2005 May 3;102 Suppl 1:6535-42. PubMed PMID: 15851677. Pubmed Central PMCID: PMC1131864. Epub 2005/04/27. eng.

[15] Manoukis NC, Powell JR, Toure MB, Sacko A, Edillo FE, Coulibaly MB, et al. A test of the chromosomal theory of ecotypic speciation in *Anopheles gambiae*. Proc Natl Acad Sci U S A. 2008 Feb 26;105(8):2940-5. PubMed PMID: 18287019. Pubmed Central PMCID: PMC2268564. Epub 2008/02/22. eng.

[16] White BJ, Cheng C, Simard F, Costantini C, Besansky NJ. Genetic association of physically unlinked islands of genomic divergence in incipient species of *Anopheles gambiae*. Mol Ecol. 2010 Mar;19(5):925-39. PubMed PMID: 20149091. Epub 2010/02/13. eng.

[17] Sinka ME, Bangs MJ, Manguin S, Coetzee M, Mbogo CM, Hemingway J, et al. The dominant *Anopheles* vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic precis. Parasit Vectors. 2010;3:117. PubMed PMID: 21129198. Pubmed Central PMCID: PMC3016360. Epub 2010/12/07. eng.

[18] Pombi M, Caputo B, Simard F, Di Deco MA, Coluzzi M, della Torre A, et al. Chromosomal plasticity and evolutionary potential in the malaria vector *Anopheles gambiae* sensu stricto: insights from three decades of rare paracentric inversions. BMC Evol Biol. 2008;8:309. PubMed PMID: 19000304. Pubmed Central PMCID: PMC2654565. Epub 2008/11/13. eng.

[19] Cornel A. *A. gambiae* s.l. Ovarian Polytene Chromosome Preparation. 2010. In: Methods in *Anopheles* Research [Internet]. CDC/ATCC/NIH/NIAID. 2nd. [242-7]. Available from: http://www.mr4.org/Portals/3/MR4_Publications/Anopheles%20Protocol%20Manual%20Second%20Ed%20v2011/2011%20Complete%20Manual%20PDF%20TOC.pdf.

[20] Dorn P, Noireau F, Krafsur ES, Lanzaro GC, Cornel AJ. Genetics of major insect vectors. In: Tibayrenc M, editor. Genetics and evolution of infectious diseases. Boston: Elsevier; 2011. p. 411-72.

[21] Touré YT, Petrarca V, Traoré SF, Coulibaly A, Maiga HM, Sankare O, et al. The distribution and inversion polymorphism of chromosomally recognized taxa of the *Anopheles gambiae* complex in Mali, West Africa. Parassitologia. 1998 Dec;40(4):477-511. PubMed PMID: 10645562. Epub 2000/01/25. eng.

[22] Coluzzi M, Sabatini A, della Torre A, Di Deco MA, Petrarca V. A polytene chromosome analysis of the *Anopheles gambiae* species complex. Science. 2002 Nov 15;298(5597):1415-8. PubMed PMID: 12364623. Epub 2002/10/05. eng.

[23] Lee Y, Cornel AJ, Meneses CR, Fofana A, Andrianarivo AG, McAbee RD, et al. Ecological and genetic relationships of the Forest-M form among chromosomal and molecular forms of the malaria vector *Anopheles gambiae* sensu stricto. Malar J. 2009;8:75. PubMed PMID: 19383163. Pubmed Central PMCID: PMC2680901. Epub 2009/04/23. eng.

[24] Cheng C, White BJ, Kamdem C, Mockaitis K, Costantini C, Hahn MW, et al. Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-resequencing approach. Genetics. 2012 Apr;190(4):1417-32. PubMed PMID: 22209907. Pubmed Central PMCID: PMC3316653. Epub 2012/01/03. eng.

[25] Coluzzi M, Sabatini A, Petrarca V, Di Deco MA. Behavioural divergences between mosquitoes with different inversion karyotypes in polymorphic populations of the *Anopheles gambiae* complex. Nature. 1977 Apr 28;266(5605):832-3. PubMed PMID: 865604. Epub 1977/04/28. eng.

[26] della Torre A, Costantini C, Besansky NJ, Caccone A, Petrarca V, Powell JR, et al. Speciation within *Anopheles gambiae*--the glass is half full. Science. 2002 Oct 4;298(5591):115-7. PubMed PMID: 12364784. Epub 2002/10/05. eng.

[27] Lee Y, Collier TC, Manoukis NC, Lozano-Fuentes S, Vallejo EE, Yan W, et al. An Individual-level Population Genomics Database for Arthropod Disease Vectors (PopI). Davis, CA2004 [updated 2012; cited 2012]. Available from: https://grassi2.ucdavis.edu/.

[28] Bryan JH, Di Deco MA, Petrarca V, Coluzzi M. Inversion polymorphism and incipient speciation in *Anopheles gambiae* s.str. in The Gambia, West Africa. Genetica. 1982;59:167-76.

[29] Coluzzi M, Petrarca V, Di Deco MA. Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae*. Bollettino di Zoologia. 1985;52(1-2):45-63.

[30] Taylor C, Touré YT, Carnahan J, Norris DE, Dolo G, Traoré SF, Edillo FE, Lanzaro GC. Gene flow among populations of the malaria vector, *Anopheles gambiae*, in Mali, West Africa. Genetics. 2001;157(2):743-50.

[31] Persiani A, Di Deco MA, Petrangeli G. [Laboratory observation of inversion polymorphisms originating from the crossing of various populations of *Anopheles gambiae* s.s]. Ann Ist Super Sanita. 1986;22(1):221-3. PubMed PMID: 3752798. Epub 1986/01/01. Osservazioni di laboratorio su polimorfismi da inversione originati da incrocitra popolazioni diverse di *Anopheles gambiae* s.s. ita.

[32] Nei M. Genetic Distance between Populations. Am Nat. 1972;106(949):283-&. PubMed PMID: WOS:A1972M475000002. English.

[33] Cianchi R, Villani F, Touré YT, Petrarca V, Bullini L. Electrophoretic study of different chromosomal forms of *Anopheles gambiae* s.s. Parassitologia. 1983;25:239-41.

[34] Lanzaro GC, Toure YT, Carnahan J, Zheng L, Dolo G, Traore S, et al. Complexities in the genetic structure of *Anopheles gambiae* populations in west Africa as revealed by microsatellite DNA analysis. Proc Natl Acad Sci U S A. 1998 Nov 24;95(24):14260-5. PubMed PMID: 9826688. Pubmed Central PMCID: PMC24361. Epub 1998/11/25. eng.

[35] Wang R, Zheng L, Toure YT, Dandekar T, Kafatos FC. When genetic distance matters: measuring genetic differentiation at microsatellite loci in whole-genome scans of recent and incipient mosquito species. Proc Natl Acad Sci U S A. 2001 Sep 11;98(19): 10769-74. PubMed PMID: 11553812. Pubmed Central PMCID: PMC58550. Epub 2001/09/13. eng.

[36] Tripet F, Dolo G, Lanzaro GC. Multilevel analyses of genetic differentiation in *Anopheles gambiae* s.s. reveal patterns of gene flow important for malaria-fighting mosquito projects. Genetics. 2005 Jan;169(1):313-24. PubMed PMID: 15677750. Pubmed Central PMCID: PMC1448890. Epub 2005/01/29. eng.

[37] Wang-Sattler R, Blandin S, Ning Y, Blass C, Dolo G, Touré YT, della Torre A, Lanzaro GC, Steinmetz LM, Kafatos FC, Zheng L. Mosaic genome architecture of the *Anopheles gambiae* species complex. PLoS One. 2007 2007 Nov 28;2(11):e1249.

[38] Coluzzi M. Spatial distribution of chromosomal inversions and speciation in Anopheline mosquitoes. Prog Clin Biol Res. 1982;96:143-53. PubMed PMID: 7178155. Epub 1982/01/01. eng.

[39] Nosil P, Funk DJ, Ortiz-Barrientos D. Divergent selection and heterogeneous genomic divergence. Mol Ecol. 2009 Feb;18(3):375-402. PubMed PMID: 19143936. Epub 2009/01/16. eng.

[40] Simard F, Ayala D, Kamdem GC, Pombi M, Etouna J, Ose K, et al. Ecological niche partitioning between *Anopheles gambiae* molecular forms in Cameroon: the ecological side of speciation. BMC Ecol. 2009;9:17. PubMed PMID: 19460146. Pubmed Central PMCID: PMC2698860. Epub 2009/05/23. eng.

[41] Neafsey DE, Lawniczak MK, Park DJ, Redmond SN, Coulibaly MB, Traore SF, et al. SNP genotyping defines complex gene-flow boundaries among African malaria vec-

tor mosquitoes. Science. 2010 Oct 22;330(6003):514-7. PubMed PMID: 20966254. Epub 2010/10/23. eng.

[42] Lawniczak MK, Emrich SJ, Holloway AK, Regier AP, Olson M, White B, et al. Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. Science. 2010 Oct 22;330(6003):512-4. PubMed PMID: 20966253. Epub 2010/10/23. eng.

[43] Coulibaly MB, Pombi M, Caputo B, Nwakanma D, Jawara M, Konate L, et al. PCR-based karyotyping of *Anopheles gambiae* inversion 2Rj identifies the BAMAKO chromosomal form. Malar J. 2007;6:133. PubMed PMID: 17908310. Pubmed Central PMCID: PMC2134931. Epub 2007/10/03. eng.

[44] White BJ, Santolamazza F, Kamau L, Pombi M, Grushko O, Mouline K, et al. Molecular karyotyping of the 2La inversion in *Anopheles gambiae*. Am J Trop Med Hyg. 2007 Feb;76(2):334-9. PubMed PMID: 17297045. Epub 2007/02/14. eng.

[45] White BJ, Cheng C, Sangare D, Lobo NF, Collins FH, Besansky NJ. The population genomics of trans-specific inversion polymorphisms in *Anopheles gambiae*. Genetics. 2009 Sep;183(1):275-88. PubMed PMID: 19581444. Pubmed Central PMCID: PMC2746151. Epub 2009/07/08. eng.

[46] Lee Y, Collier TC, Sanford MR, Marsden CD, Fofana A, Cornel AJ, et al. Chromosome Inversions, Genomic Differentiation and Speciation in the African Malaria Mosquito *Anopheles gambiae*. PLoS One. 2013.

[47] Favia G, Lanfrancotti A, Spanos L, Siden-Kiamos I, Louis C. Molecular characterization of ribosomal DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae* s.s. Insect Mol Biol. 2001 Feb;10(1):19-23. PubMed PMID: 11240633. Epub 2001/03/10. eng.

[48] Favia G, della Torre A, Bagayoko M, Lanfrancotti A, Sagnon N, Toure YT, et al. Molecular identification of sympatric chromosomal forms of *Anopheles gambiae* and further evidence of their reproductive isolation. Insect Mol Biol. 1997 Nov;6(4):377-83. PubMed PMID: 9359579. Epub 1997/11/14. eng.

[49] della Torre A, Tu Z, Petrarca V. On the distribution and genetic differentiation of *Anopheles gambiae* s.s. molecular forms. Insect Biochem Mol Biol. 2005 Jul;35(7):755-69. PubMed PMID: 15894192. Epub 2005/05/17. eng.

[50] Costantini C, Ayala D, Guelbeogo WM, Pombi M, Some CY, Bassole IH, et al. Living at the edge: biogeographic patterns of habitat segregation conform to speciation by niche expansion in *Anopheles gambiae*. BMC Ecol. 2009;9:16. PubMed PMID: 19460144. Pubmed Central PMCID: PMC2702294. Epub 2009/05/23. eng.

[51] Fanello C, Santolamazza F, della Torre A. Simultaneous identification of species and molecular forms of the *Anopheles gambiae* complex by PCR-RFLP. Med Vet Entomol. 2002 Dec;16(4):461-4. PubMed PMID: 12510902. Epub 2003/01/04. eng.

[52] Santolamazza F, Caputo B, Calzetta M, Vicente JL, Mancini E, Petrarca V, et al. Comparative analyses reveal discrepancies among results of commonly used methods for *Anopheles gambiae* molecular form identification. Malar J. 2011;10:215. PubMed PMID: 21810255. Pubmed Central PMCID: PMC3170251. Epub 2011/08/04. eng.

[53] Masendu HT, Hunt RH, Govere J, Brooke BD, Awolola TS, Coetzee M. The sympatric occurrence of two molecular forms of the malaria vector *Anopheles gambiae* Giles sensu stricto in Kanyemba, in the Zambezi Valley, Zimbabwe. Trans R Soc Trop Med Hyg. 2004 Jul;98(7):393-6. PubMed PMID: 15138074. Epub 2004/05/13. eng.

[54] della Torre A, Fanello C, Akogbeto M, Dossou-yovo J, Favia G, Petrarca V, et al. Molecular evidence of incipient speciation within *Anopheles gambiae* s.s. in West Africa. Insect Mol Biol. 2001 Feb;10(1):9-18. PubMed PMID: 11240632. Epub 2001/03/10. eng.

[55] Tripet F, Toure YT, Taylor CE, Norris DE, Dolo G, Lanzaro GC. DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. Mol Ecol. 2001 Jul;10(7):1725-32. PubMed PMID: 11472539. Epub 2001/07/27. eng.

[56] Wondji C, Frederic S, Petrarca V, Etang J, Santolamazza F, Della Torre A, et al. Species and populations of the *Anopheles gambiae* complex in Cameroon with special emphasis on chromosomal and molecular forms of Anopheles gambiae s.s. J Med Entomol. 2005 Nov;42(6):998-1005. PubMed PMID: 16465741. Epub 2006/02/10. eng.

[57] Caputo B, Nwakanma D, Jawara M, Adiamoh M, Dia I, Konate L, et al. *Anopheles gambiae* complex along The Gambia river, with particular reference to the molecular forms of *An. gambiae* s.s. Malar J. 2008;7:182. PubMed PMID: 18803885. Pubmed Central PMCID: PMC2569043. Epub 2008/09/23. eng.

[58] Oliveira E, Salgueiro P, Palsson K, Vicente JL, Arez AP, Jaenson TG, et al. High levels of hybridization between molecular forms of *Anopheles gambiae* from Guinea Bissau. J Med Entomol. 2008 Nov;45(6):1057-63. PubMed PMID: 19058629. Epub 2008/12/09. eng.

[59] Marsden CD, Lee Y, Nieman CC, Sanford MR, Dinis J, Martins C, et al. Asymmetric introgression between the M and S forms of the malaria vector, *Anopheles gambiae*, maintains divergence despite extensive hybridization. Mol Ecol. 2011 Dec;20(23):4983-94. PubMed PMID: 22059383. Pubmed Central PMCID: PMC3222736. Epub 2011/11/09. eng.

[60] Riehle MM, Guelbeogo WM, Gneme A, Eiglmeier K, Holm I, Bischoff E, et al. A cryptic subgroup of *Anopheles gambiae* is highly susceptible to human malaria parasites. Science. 2011 Feb 4;331(6017):596-8. PubMed PMID: 21292978. Pubmed Central PMCID: PMC3065189. Epub 2011/02/05. eng.

[61] Crawford JE, Bischoff E, Garnier T, Gneme A, Eiglmeier K, Holm I, et al. Evidence for Population-Specific Positive Selection on Immune Genes of *Anopheles gambiae*. G3

(Bethesda). 2012 Dec;2(12):1505-19. PubMed PMID: 23275874. Pubmed Central PMCID: PMC3516473. Epub 2013/01/01. eng.

[62] Caputo B, Santolamazza F, Vicente JL, Nwakanma DC, Jawara M, Palsson K, et al. The "far-west" of *Anopheles gambiae* molecular forms. PLoS One. 2011;6(2):e16415. PubMed PMID: 21347223. Pubmed Central PMCID: PMC3039643. Epub 2011/02/25. eng.

[63] Diabate A, Dabire RK, Millogo N, Lehmann T. Evaluating the effect of postmating isolation between molecular forms of *Anopheles gambiae* (Diptera: Culicidae). J Med Entomol. 2007 Jan;44(1):60-4. PubMed PMID: 17294921. Epub 2007/02/14. eng.

[64] Di Deco MA, Petrarca V, Villani F, Coluzzi M. Polimorfismo cromosomico d inversioni paracentriche ed eccesso degli eterocariotipi in ceppi di *Anopheles allevati* in laboratorio. Parassitologia. 1980;22:304-7.

[65] Diabate A, Dabire RK, Kengne P, Brengues C, Baldet T, Ouari A, et al. Mixed swarms of the molecular M and S forms of *Anopheles gambiae* (Diptera: Culicidae) in sympatric area from Burkina Faso. J Med Entomol. 2006 May;43(3):480-3. PubMed PMID: 16739404. Epub 2006/06/03. eng.

[66] Lehmann T, Diabate A. The molecular forms of *Anopheles gambiae*: a phenotypic perspective. Infect Genet Evol. 2008 Sep;8(5):737-46. PubMed PMID: 18640289. Pubmed Central PMCID: PMC2731232. Epub 2008/07/22. eng.

[67] Tripet F, Wright J, Cornel A, Fofana A, McAbee R, Meneses C, et al. Longitudinal survey of knockdown resistance to pyrethroid (kdr) in Mali, West Africa, and evidence of its emergence in the Bamako form of *Anopheles gambiae* s.s. Am J Trop Med Hyg. 2007 Jan;76(1):81-7. PubMed PMID: 17255234. Epub 2007/01/27. eng.

[68] Lee Y, Meneses CR, Fofana A, Lanzaro GC. Desiccation resistance among subpopulations of *Anopheles gambiae* s.s. from Selinkenyi, Mali. J Med Entomol. 2009 Mar;46(2): 316-20. PubMed PMID: 19351082. Epub 2009/04/09. eng.

[69] Gimonneau G, Pombi M, Choisy M, Morand S, Dabire RK, Simard F. Larval habitat segregation between the molecular forms of the mosquito *Anopheles gambiae* in a rice field area of Burkina Faso, West Africa. Med Vet Entomol. 2012 Mar;26(1):9-17. PubMed PMID: 21501199. Pubmed Central PMCID: PMC3140611. Epub 2011/04/20. eng.

[70] Sanford MR, Demirci B, Marsden CD, Lee Y, Cornel AJ, Lanzaro GC. Morphological differentiation may mediate mate-choice between incipient species of *Anopheles gambiae* s.s. PLoS One. 2011;6(11):e27920. PubMed PMID: 22132169. Pubmed Central PMCID: PMC3221689. Epub 2011/12/02. eng.

[71] Diabate A, Dao A, Yaro AS, Adamou A, Gonzalez R, Manoukis NC, et al. Spatial swarm segregation and reproductive isolation between the molecular forms of

Anopheles gambiae. Proc Biol Sci. 2009 Dec 7;276(1676):4215-22. PubMed PMID: 19734189. Pubmed Central PMCID: PMC2821344. Epub 2009/09/08. eng.

[72] Manoukis NC, Diabate A, Abdoulaye A, Diallo M, Dao A, Yaro AS, et al. Structure and dynamics of male swarms of *Anopheles gambiae*. J Med Entomol. 2009 Mar;46(2): 227-35. PubMed PMID: 19351073. Pubmed Central PMCID: PMC2680012. Epub 2009/04/09. eng.

[73] Lehmann T, Licht M, Elissa N, Maega BT, Chimumbwa JM, Watsenga FT, et al. Population structure of *Anopheles gambiae* in Africa. J Hered. 2003 Mar-Apr;94(2):133-47. PubMed PMID: 12721225. Epub 2003/05/02. eng.

[74] Slotman MA, Reimer LJ, Thiemann T, Dolo G, Fondjo E, Lanzaro GC. Reduced recombination rate and genetic differentiation between the M and S forms of *Anopheles gambiae* s.s. Genetics. 2006 Dec;174(4):2081-93. PubMed PMID: 17057242. Pubmed Central PMCID: PMC1698612. Epub 2006/10/24. eng.

[75] Stump AD, Fitzpatrick MC, Lobo NF, Traore S, Sagnon N, Costantini C, et al. Centromere-proximal differentiation and speciation in *Anopheles gambiae*. Proc Natl Acad Sci U S A. 2005 Nov 1;102(44):15930-5. PubMed PMID: 16247019. Pubmed Central PMCID: PMC1276105. Epub 2005/10/26. eng.

[76] Turner TL, Hahn MW, Nuzhdin SV. Genomic islands of speciation in *Anopheles gambiae*. PLoS Biol. 2005 Sep;3(9):e285. PubMed PMID: 16076241. Pubmed Central PMCID: PMC1182689. Epub 2005/08/04. eng.

[77] Turner TL, Hahn MW. Locus- and population-specific selection and differentiation between incipient species of *Anopheles gambiae*. Mol Biol Evol. 2007 Sep;24(9):2132-8. PubMed PMID: 17636041. Epub 2007/07/20. eng.

[78] Wu CI. The genic view of the process of speciation. J Evol Biol. 2001;14:851-65.

[79] Wu CI, Ting CT. Genes and speciation. Nat Rev Genet. 2004 Feb;5(2):114-22. PubMed PMID: 14735122. Epub 2004/01/22. eng.

[80] Noor MA, Grams KL, Bertucci LA, Reiland J. Chromosomal inversions and the reproductive isolation of species. Proc Natl Acad Sci U S A. 2001 Oct 9;98(21):12084-8. PubMed PMID: 11593019. Pubmed Central PMCID: PMC59771. Epub 2001/10/11. eng.

[81] Rieseberg LH. Chromosomal rearrangements and speciation. Trends Ecol Evol. 2001 Jul 1;16(7):351-8. PubMed PMID: 11403867. Epub 2001/06/14. Eng.

[82] Turner TL, Hahn MW. Genomic islands of speciation or genomic islands and speciation? Mol Ecol. 2010 Mar;19(5):848-50. PubMed PMID: 20456221. Epub 2010/05/12. eng.

[83] Hahn MW, White BJ, Muir CD, Besansky NJ. No evidence for biased co-transmission of speciation islands in *Anopheles gambiae*. Philos Trans R Soc Lond B Biol Sci. 2012

Feb 5;367(1587):374-84. PubMed PMID: 22201167. Pubmed Central PMCID: PMC3233708. Epub 2011/12/28. eng.

[84] Slotman MA, Tripet F, Cornel AJ, Meneses CR, Lee Y, Reimer LJ, Thiemann TC, Fondjo E, Fofana A, Traoré SF, Lanzaro GC. Evidence for subdivision within the M molecular form of *Anopheles gambiae*. Mol Ecol. 2007;16(3):639-49.

IntechOpen

IntechOpen