Infection, Genetics and Evolution xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Discussion

The molecular forms of *Anopheles gambiae*: A phenotypic perspective

Tovi Lehmann *, Abdoulaye Diabate

Laboratory of Malaria and Vector Research, NIAID, NIH, MS 8132, 12735 Twinbrook Parkway, Rockville, MD, USA

ARTICLE INFO

Article history:
Received 13 February 2008
Received in revised form 11 June 2008
Accepted 12 June 2008
Available online xxx

Keywords:
Anopheles gambiae
Malaria
Mating behavior
Predation
Reproductive isolation
Divergent selection
Ecological speciation
Swarm

ABSTRACT

The African malaria mosquito *Anopheles gambiae* is undergoing speciation, being split into the M and S molecular forms. Speciation is the main process promoting biological diversity, thus, new vector species might complicate disease transmission. Genetic differentiation between the molecular forms has been extensively studied, but phenotypic differences between them, the evolutionary forces that generated divergence, and the mechanisms that maintain their genetic isolation have only recently been addressed. Here, we review recent studies suggesting that selection mediated by larval predation and competition promoted divergence between temporary and permanent freshwater habitats. These differences explain the sharp discontinuity in distribution of the molecular forms between rice fields and surrounding savanna, but they can also explain the concurrent cline between humid and arid environments due to the dependence on permanent habitats in the latter. Although less pronounced, differences in adult body size, reproductive output, and longevity also suggest that the molecular forms have adapted to distinct niches. Reproductive isolation between the molecular forms is achieved by spatial swarm segregation, although within-swarm mate recognition appears to play a role in certain locations. The implications of these results to disease transmission and control are discussed and many of the gaps in our understanding are highlighted.

Published by Elsevier B.V.

1. Introduction

The principal African malaria vector, *Anopheles gambiae* sensu stricto is one of seven morphologically undistinguishable sibling species. The definitive evidence for the existence of the *A. gambiae* complex was obtained by crossing experiments revealing complete or partial sterility in the hybrids (Davidson, 1964), although, the taxonomic status of the salty-water *vs.* fresh water—"varieties" has been debated earlier (e.g., by Ribbands, 1944). Crossing experiments led to the recognition of three fresh water species: *A. gambiae* s.s., *A. arabiensis*, *A. quadriannulatus*, and three brackish water species: *A. bwambae, A. melas* and *A. merus*, and later to the division of *A. quadriannulatus* into two species (Hunt et al., 1998). Fixed chromosomal inversions between these species (e.g., Coluzzi and Sabatini, 1967) facilitated simpler identification of field specimens and led to important genetic and ecological discoveries. The phylogenetic relationship among the members of the complex is obscured by introgression. For example, the similarity between the most important vectors *A. gambiae* s.s. and *A. arabiensis* may be a result of extensive introgression rather than closer phylogenetic relationship (Krzywinski and Besansky, 2003).

A. gambiae s.s. is undergoing speciation, being split into two "molecular forms", currently named 'M' and 'S'. Speciation is the main process promoting biological diversity and in the context of public health it increases epidemiological complexity. New species of pathogens and vectors might change disease manifestations and transmission patterns if they differ in traits affecting pathogen virulence or vectorial capacity. The study of closely related species and incipient species has been central for the current understanding of speciation (Coyne and Orr, 1997; Grant and Grant, 1979; Mallet, 2006; Wu and Ting, 2004), but many of its aspects remain unclear if not enigmatic. The seven sibling and two incipient species of the *A. gambiae* complex provide an excellent opportunity to extend basic understanding of speciation as well as provide essential public health information.

Genetic differentiation between the molecular forms and its peculiar distribution across the genome has been extensively studied and several good reviews cover the main findings (e.g., Black and Lanzaro, 2001; della Torre et al., 2002; Krzywinski and Besansky, 2003), but the evolutionary forces that generated divergence, the particulars of the accompanying phenotypic divergence, and the mechanisms that maintain genetic isolation between them received little attention. Thoughtful speculation

* Corresponding author. Tel.: +1 301 451 1059; fax: +1 301 480 2038.
E-mail address: tlehmann@niaid.nih.gov (T. Lehmann).

1567-1348/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.meegid.2008.06.003

about the role of human generated habitats in promoting speciation (Coluzzi et al., 2002) is frequently cited, but the ecological conditions that promote speciation in *A. gambiae* have only recently been subjected to experimental study. Here, we review advances in understanding the status of the molecular forms, focusing on the underlying ecological conditions that promoted divergence and speciation. Recent studies revealed phenotypic divergence between the molecular forms across life stages. Larval adaptations to exploit temporary vs. 'permanent' freshwater habitats account for discontinuities in distribution of the molecular forms (Diabate et al., 2008) and provide evidence for ecological speciation (e.g., Schluter, 2001). Divergence in adult mating behavior involves spatial swarm segregation and within-swarm mate recognition but they appear to play different roles in promoting reproductive isolation in Burkina Faso and Mali.

2. The recognition of the molecular forms

Early studies of *A. gambiae* in West Africa based on chromosomal inversions revealed deficits of heterozygotes (with respect to corresponding Hardy Weinberg expectations) in a number of inversions. Based on these data, five partly isolated populations were defined and named Forest, Savanna, Bamako, Mopti, and Bissau chromosomal forms (Bryan et al., 1982; Coluzzi et al., 1985, 1979; Toure et al., 1998). Inversion frequencies were strongly correlated with aridity on spatial and temporal scales (Coluzzi et al., 1985, 1979; Toure et al., 1994), indicating that selection had a key role in shaping inversion frequencies, hence it confounds assessment of genetic isolation between forms. No marked differentiation was found between the chromosomal forms using allozymes (Cianchi et al., 1983) and microsatellites outside inversions on chromosome II (Lanzaro et al., 1998), but fixed differences were found in the rDNA intergenic spacer located on the X chromosome separating Mopti from both Savanna and Bamako chromosomal forms in Mali (Favia et al., 1997). The two rDNA haplotypes were used to define the molecular forms (della Torre et al., 2001; Favia et al., 1997; Favia et al., 2001).

Subsequent studies using multiple gene sequences (della Torre et al., 2001; Favia et al., 2001; Gentile et al., 2001; Mukabayire et al., 2001), microsatellites (Lehmann et al., 2003; Wang et al., 2001; Wondji et al., 2002), and site occupancy of transposable elements (della Torre et al., 2005) revealed a consistent pattern: (1) high differentiation ($Fst > 0.2$) loci were few and clustered into two genomic regions: near the centromere of chromosome X where the diagnostic rDNA polymorphism resides (Lehmann et al., 2003; Stump et al., 2005; Wang et al., 2001) and the centromere of chromosome II near the sodium channel gene (Chandre et al., 1999), whereas across most of the genome, differentiation was slight or none ($Fst < 0.05$); (2) differentiation was detected only between two populations rather than five. In Mali and Burkina Faso, these groups corresponded to the Mopti vs. Savanna and Bamako (Favia et al., 1997; Favia and Louis, 1999), hence the terms M and S molecular forms. This correspondence, however, did not hold for specimens from adjacent countries (della Torre et al., 2001), for example the forest chromosomal form in Central Africa was divided into the sympatric molecular forms and no M/S hybrids were found (Lehmann et al., 2003; Wondji et al., 2002). The most recent addition to these results was the finding of differentiation ($Fst = 0.045$) between M molecular form from Mali (M-Mopti) and Cameroon (M-Forest), which may reflect a secondary subdivision in the M form (Slotman et al., 2007). Notably, no postmating isolation was detected by laboratory crosses of colonies representing the chromosomal forms (Di Deco et al., 1980; Persiani et al., 1986) nor in crosses of F1s from naturally collected mothers representing the molecular forms from Burkina Faso (Diabate et al., 2007). Nevertheless, incomplete but strong assortative mating was found between sympatric populations of the molecular forms in Mali (Tripet et al., 2001), where cross mating (1.2%) was substantially lower than expected by chance (17%).

The interpretation of this pattern has been contentious. Many authors considered that it represents recent reproductive isolation between two incipient species (e.g., della Torre et al., 2002, 2001; Favia et al., 2001; Mukabayire et al., 2001). Others argued that reproductive isolation, even if recent, cannot account for the sharp heterogeneity between parts of the genome, and emphasized that it is selection operating on few genomic regions in face of gene flow that shapes this pattern (e.g., Lanzaro et al., 1998; Lehmann et al., 2003). If selection operates on few genomic regions, the forms cannot represent incipient species, because most of their gene pools evolve together. How many loci show high differentiation and how they are distributed throughout the genome was not completely clear. Recently, these issues were addressed using whole genome micro-array that surprisingly affirmed that high differentiation loci are concentrated only in the two previously known (and maybe a third) miniscule regions (above), which together encompass only ~1% of the genome (Turner et al., 2005).

Underlying these opposing views were basic notions of speciation process (Schluter, 2001; Wu, 2001; Wu and Ting, 2004). Conventional models of speciation required effectively complete barrier to gene flow from the beginning of the process as in allopatric speciations models, resulting in genome wide homogenous differentiation, increasing over time. Recent models considered situations with incomplete barriers to gene flow, in which divergent selection operating on a few loci can gradually increase differentiation between populations until gene flow essentially ceases, specifically predicting heterogeneity across the genome (Wu, 2001; Wu and Ting, 2004). In addition to accommodating the apparently contradictory findings (above), direct evidence for this model showing selection signatures on the high differentiation "speciation islands" (tiny chromosomal regions encompassing less than 1% of the genome) in *A. gambiae* was obtained (Turner and Hahn, 2007). These theoretical and empirical advances have mostly resolved the controversy and the molecular forms are now perceived as incipient species by many, if not all medical entomologists.

2.1. What remains of the chromosomal forms?

Inversions have been considered the basis for the speciation in *A. gambiae* s.s. (Coluzzi et al., 1985, 2002). Most studies, however, were conducted near Mali and Burkina Faso where the chromosomal forms perfectly corresponded to the molecular forms. The lack of correspondence between molecular and chromosomal forms in other areas coupled with the near absence of M/S hybrids questioned the significance of the latter (della Torre et al., 2001). Until complete genome comparison is done, it remains possible that new speciation islands will be found between Bamako and Savanna chromosomal forms (Manoukis et al., 2008) despite extensive genomic surveys (Slotman et al., 2006). However, disconnected from the molecular forms, the chromosomal forms do not correspond to genetically distinct populations based on currently available data. Instead, they appear to represent conspecific populations subject to different selection pressures with different frequencies of inversion "alleles". Early studies revealed a latitudinal cline in inversion frequencies with aridity combined with corresponding seasonal fluctuations in inversion frequencies (Coluzzi et al., 1985; Toure et al., 1994, 1998), providing solid ground for the role of selection on inversions. These selection pressures are at least partially distinct from those causing the divergence of the molecular forms (della Torre et al.,

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003

2002). In arid pre-desert areas of Angola, the Forest chromosomal form was found despite the absence of inversions, such as 2La and 2Rb that are common in the similarly arid Sahel (Calzetta et al., 2008). Notably, these Forest mosquitoes were solely of the M molecular form, suggesting that adaptation to arid environment is not only conferred by inversions.

The small geographical range of Bamako and Bissau chromosomal forms which encompass only hundreds of km^2 (Fig. 1) raises the question of what are the strong ecological factors promoting such divergence (at least in terms of the frequency of certain inversions) and how these populations could diverge and maintain their distinct inversion combinations in the face of extensive gene flow from sympatric Savanna and Mopti forms. Similar problems pertain to *A. bwambae*, residing in the Similiki Forest in Uganda that exploits mineral springs as its larval sites. Nevertheless, this does not imply that inversions have not facilitated speciation (Ayala and Coluzzi, 2005).

3. Phenotypic divergence between the molecular forms

Divergent selection is fundamental to several speciation models as the underlying cause of differential fitness of genotypes in different environments which directly or indirectly promotes reproductive isolation (Schluter, 2001). Notably, phenotypic divergence alone may have epidemiological implications; if it changes vectorial capacity, vector's distribution range, or its ability to withstand means of control. Besides studies on composition of the molecular forms in various locations and on insecticide resistance, few studies focused on phenotypic differences between the molecular forms. Uncovering the ecological factors that promote divergent selection between the molecular forms of *A. gambiae* is a complicated task. The larva (and egg) of *A. gambiae* experiences distinct ecological settings from the adult, and accordingly, such assessment requires examination of phenotypic variation between the molecular forms at all life stages (Table 1). Most studies comparing the molecular forms have not fully accommodated variation between populations within-form. Phenotypic variation may depend on the karyotype (see above), thus differences between Forest-M and Mopti populations would be larger than between two Forest-M populations. Most studies have compared M and S populations in the savanna, which correspond to the Mopti (M) vs. Savanna (S); although in Mali the S form could contain an unknown fraction of Bamako. No studies compared S populations from West and East Africa, so generalization should be considered cautiously.

3.1. Geographical distribution patterns

Understanding the evolution of the molecular forms must incorporate their distribution range. With one possible exception finding two M form specimens in East Africa (Masendu et al., 2004), there is no indication for change in these boundaries, hence we assume their ranges are rather stable (Fig. 1). In this schematic map portraying the geographical relations between the molecular and chromosomal forms, we pooled the intergrading Savanna (which is found throughout the savanna belt across the continent) and Forest (which covers the narrower mostly coastal humid rainforest belt from Liberia to Democratic Republic of Congo) chromosomal forms for simplification. Three important features are apparent from examination of this schematic map: (i) extensive overlap in geographical range, (ii) boundaries incompatible with known bio-geographical factors, and (iii) substantial variation in range's size.

Over 90% of the range of the M form overlaps with that of the S form, suggesting considerable phenotypic overlap between them. Nonetheless, the narrow semi-desert belt inhabited only by the M

Table 1
Summary of phenotypic comparisons between the molecular forms

Trait	M	S	CO[a]	Reference[b]
Adaptation to larval & adult habitat				
Geographical range	West + Central	Continental	H	(della Torre et al., 2005)
Arid-mesic gradient (spatial + season)	Dry	Wet	H	(Coluzzi et al., 1979, 1985)
Adaptation to larval habitat				
Egg hatch timing & responses	Fast	Fast[c]	L	(Yaro et al., 2006a)
Egg desiccation tolerance	Low	Low[c]	H	Dao et al. unpublished
Larval habitat (preferred)	Permanent (rice)	Temporary (puddle)[c]	H	(Diabate et al., 2002, 2005; Robert et al., 1988; Toure et al., 1998)
Larva predator avoidance	Higher	Lower[c]	M	(Diabate et al., 2008)
Larva competitiveness (no predators)	Lower	Higher[c]	M	(Diabate et al., 2005)
Larval developmental time	Slower	Faster[c]	M	(Diabate et al., 2008)
Adaptation to adult environment				
Longevity (adult)	Longer	Shorter[c]	L	Dao et al.: unpublished
Body size (adult)	Larger	Smaller[c]	L	(Yaro et al., 2006b)
Reproductive output	Larger	Smaller[c]	L	(Yaro et al., 2006b)
Anthropophily	High	High[c]	L	(Wondji et al., 2005a)
Endophily	High	High	L	(Wondji et al., 2005a)
Plasmodium susceptibility	High	High	M	(Wondji et al., 2005b; Yaro et al., 2006b)
Insecticide resistance (kdr)	Low	High	H	(Chandre et al., 1999; Tripet et al., 2007; Yawson et al., 2004)
Mating Behavior				
Flight tone	492 Herz	493 Herz[c]	M	(Tripet et al., 2004)
Swarm landmark type	High contrast	Bare ground[c]	H	(Diabate et al. unpublished)
Indoor mating	Low frequency	None[c]	M	(Dao et al., 2008)

Traits in which only minimal differences were found between forms are colored in gray.

[a] Confidence level: high (H), moderate (M), and low (L) refers to the uncertainty in the generality of the result due to possible confounding factor, e.g., variation due to within form inversion karyotype, locality, and time. Repeated patterns in independent studies were considered as highly confident, whereas low confidence score was assigned to a single study on a single population.

[b] Diabate et al., 2002; Robert et al., 1988; Wondji et al., 2005a,b; Tripet et al., 2007; Yawson et al., 2004; Tripet et al., 2004Studies comparing the Mopti and either Savanna or Bamako chromosomal forms in Mali and Burkina Faso indirectly compared the molecular form because of the complete correspondence between these entities, hence were cited here (see text). Representative references are listed for traits with many references e.g., insecticide resistance.

[c] Comparisons include only West African S populations and may not apply for East African populations (see text).

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003

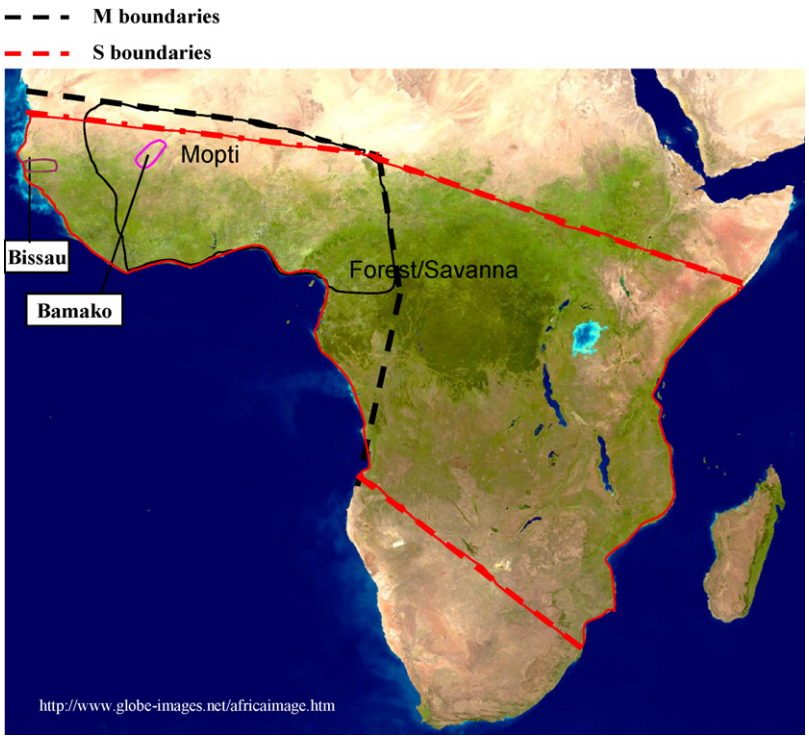

Fig. 1. A map showing schematic distribution of the molecular and chromosomal forms overlayed on a satellite image of the continent revealing vegetation pattern.

form reflects its greater aridity tolerance. Even in areas where both forms coexist, the M form predominates in arid habitats at higher latitudes (annual precipitations <1000 mm and evapotranspiration >2000 mm) whilst the S form predominates in humid lower latitudes (annual precipitations >1000 mm and evapotranspiration <1800 mm) (Bayoh et al., 2001; della Torre et al., 2001, 2005). This variation with aridity corresponds to seasonal fluctuations in form composition in the same locale (Toure et al., 1998). This pattern was complicated, however, by finding both molecular forms in sympatry in the humid forest belt (della Torre et al., 2005; Wondji et al., 2002).

The eastern boundary of the M form is puzzling because it does not coincide with known eco-geographic or biotic zones. What has allowed the S form to extend across West and East Africa but blocked the M form is a mystery and deserves focused research. Notably, in East Africa *A. arabiensis* occupies what appears to be the same niche the M form fills in West Africa. Although *A. arabiensis* extends beyond the eastern boundary of the M form into West Africa, comparing its abundance and ecology with that of the M form across this boundary might provide insights into the nature of this border. *A. arabiensis* exhibits a higher degree of zoophily in East Africa than in West and Central Africa and is conspicuously missing (or rare) from rice fields in West Africa where the M form predominates. A more detailed comparison of *A. arabiensis* ecology with that of the M form might reveal evidence of character displacement in West Africa, where the forms are sympatric. Presumably the similarities in larval ecology of *A. arabiensis* in East Africa and the M form in West Africa represent convergent evolution although the possibility of introgression cannot be ruled out (Besansky et al., 2003, 1997; Coluzzi et al., 2002; Donnelly et al., 2004). Nevertheless, the stability of the boundary of the M form should be determined, because if it advances outwards, the limited time since divergence may better explain the M form's current distribution than the considerations above.

3.2. Variation in traits of the egg and larva

The most remarkable discontinuity in the distribution of the molecular forms is displayed between rice field areas where the M form is found exclusively and surrounding savanna where the S form predominates, at least during the wet season (Table 1). The form composition flips from one extreme to the other, along few kilometers, indicating strong selective force related to larval habitats.

To evaluate adaptive differences between the aquatic stages of the molecular forms, observational and experimental studies were conducted on eggs and larvae (Table 1). Minor differences were detected in the time to hatch between eggs of the molecular forms (Yaro et al., 2006a), and no differences were found in the egg desiccation tolerance (Dao, Yaro, Adamou, and Lehmann unpublished), suggesting that eggs of both forms experience similar conditions. Larval surveys in Mali, where the molecular forms are sympatric, revealed that both forms shared many larval sites without clear segregation (Edillo et al., 2002), very similar to findings in East Africa on *A. gambiae* and *A. arabiensis* (Gimnig et al., 2001). The co-occurrence of forms in many larval sites does not imply that they have the same ability to exploit these habitats. Experimental field studies, using transplantation cages into which first instar larvae of one or both forms were placed revealed that the S form outperformed the M form in both temporary puddles as well as rice fields (Diabate et al., 2005). Comparing developmental success of the molecular forms in mixed and single form transplantation cages showed the S form outcompeted the M form in puddles (inconclusive results were obtained in rice fields). Subsequent studies using the same approach were directed to assess the role of larval predators in mediating habitat segregation (rice fields vs. puddles) between the molecular forms (Diabate et al., 2008). These studies revealed that developmental success of the molecular forms in the different larval habitats was dependent on the presence of predators: The success of the M form was higher

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003

Table 2
Effect of predators in different larval habitats on developmental success of the molecular forms (adapted based on Diabate et al., 2008)

Habitat	Predator	M form	S form	χ^2/P
Puddles	Absent	46.3% (857)[a]	53.7% (993)	11.65/0.0006
Puddles	Present	56.2% (199)	43.8% (155)	
Rice fields	Absent	55.1% (576)	44.9% (470)	2.35/0.124
Rice fields	Present	59.6% (226)	40.4% (153)	
Total (Pooled)	Absent	49.5% (1433)	50.5% (1463)	16.9/0.0001
	Present	58% (425)	42% (308)	12.4/0.0004[b]

[a] Parenthesis indicates number of adults emerged.
[b] Stratified analysis of form by predation controlling for habitat using Cochran–Mantel–Haenszel test.

than that of the S form in both habitats under predator pressure (Table 2). This study also confirmed that predators were much more abundant in rice fields than in puddles, as previously documented for permanent vs. temporary habitats (Williams, 2006). Larvae of the S form developed faster than those of the M form in both permanent and temporary habitats, possibly as a response to the higher risk of desiccation in temporary larval habitats it typically occupies (Diabate et al., 2008). Despite its slower developmental time forcing it to endure larval predators longer, the M form possesses a superior predator avoidance ability, the nature of which remains to be identified. Mosquitoes exhibit various behavioral responses to avoid larval predators (Blaustein et al., 2004; Kesavaraju et al., 2007b). Moreover, a negative relationship between predator avoidance capacity and competitive ability was detected in container mosquitoes *Aedes albopoictus* and *Ochlerotatus triseriatus* (Kesavaraju et al., 2007a). The results above provide the first empirical evidence for specific adaptive differences between the molecular forms and stress the role of larval predation as one of the selective forces contributing to their divergence. These studies explain the habitat segregation between rice fields and surrounding savanna based on the superior competitive power of the S form larvae in temporary habitats with low predation pressure, whereas the superior predator avoidance of the M form allows it to dominate in predator-rich larval habitats. Because larval habitats in arid climate tend to be more permanent (as rain is too rare to sustain populations), the capacity of the M form to exploit such larval habitat explains its extended distribution into dryer environment. Hence, the aquatic larva, rather than the terrestrial adult appears to possess the adaptation conferring aridity tolerance in the M form.

These results are consistent with previous studies showing that the length of hydroperiod in freshwater bodies mediates divergent selection in a number of species (Williams, 2006). Predation and competition also mediated divergent selection resulting in adaptive radiation in *Timema* stick insects (Nosil and Crespi, 2006a,b) and indirectly, predation promotes premating isolation (Nosil and Crespi, 2004) in walking-sticks. Insecticide resistance in *Culex pipiens* increased susceptibility to predators (Berticat et al., 2004). Because the frequency of the *kdr* resistance allele is high in the S form but low in the M form (Table 1), its possible contribution to the results above should be evaluated.

3.3. Life history and vectorial capacity traits in adults

Due to their direct impact on fitness, divergence in life history traits may accompany speciation reflecting the different conditions of the niches occupied by each species (McKinnon et al., 2004). Notably, life history traits such as longevity and reproductive output strongly affect vectorial capacity. Egg batch size of the M form was larger than that of the S form in accordance with the M form larger body size (Yaro et al., 2006b). No difference was found in the relationship between egg batch size and body size between the forms. The egg protein content of an M form was slightly higher than that of the S form, indicating that its total reproductive output is correspondingly higher than that of the S form (Yaro et al., 2006b). This could be a response to the lower overall productivity of more permanent sites (Diabate et al., 2008, 2005; Diuk-Wasser et al., 2005).

In an independent study, longevity of F1 virgin females, representing offspring of wild collected females that were raised in the laboratory under uniform conditions was longer in the M molecular form across localities (Dao, Kassogue, Adamou, Yaro, and Lehmann unpublished, Fig. 2). The effect of body size on longevity was not significant ($P > 0.45$). The significant difference between forms, treating populations as fixed effects could not be generalized to "any" M and S population if both population and the interaction between form and population were considered as random effects ($P > 0.28$) because of the large variation between populations (especially between the M populations from Niono and Dongebougou; Fig. 2). Such variation in longevity between the forms could be of epidemiological importance because slight changes in longevity greatly change vectorial capacity. However, infection prevalence with *Plasmodium falciparum* was similar between the forms (Table 1). This was also the case for anthropophily and endophily (Table 1). Because the relative contribution of environmental variation to these traits is very large (Koella and Boete, 2002; Lambrechts et al., 2006; Lehmann et al., 2006; Schwartz and Koella, 2002), independent comparisons of these traits in multiple settings will be valuable.

Despite the inability to identify the underlying ecological factor and selective pressure operating on adults, these results suggest that life history traits including body size, reproductive output, and longevity have diverged between forms as a result of their exposure to different environments. These trends would result in a

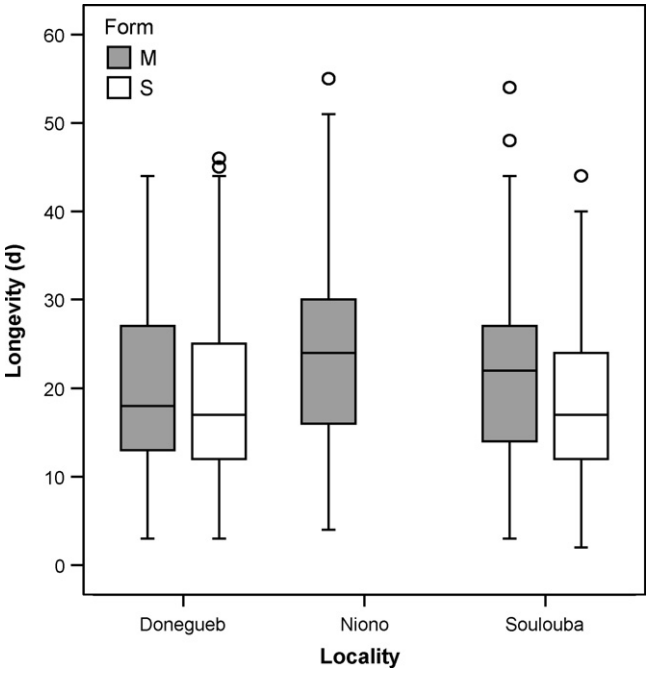

Fig. 2. Longevity of virgin F1 females of the molecular forms from Mali (fed on 5% sucrose). The forms were sympatric in Dongebougou and Soulouba (100 km apart) but only the M form was found in Niono (>300 km from other sites). In box-whisker plot, the box extends between the 25th and the 75th percentile, i.e., across one inter-quartile range (IQR), and the whiskers extend up to the most extreme value, but not beyond 1.5 times the IQR. Outlier values located 1.5–3 IQR from the median are shown as "○". Number of F1 females included in each box was 264–274.

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003

higher vectorial capacity of the M form. Contrary to this prediction, malaria transmission in villages surrounded by rice fields where the M form predominates is low (Diuk-Wasser et al., 2005; Manoukis et al., 2006), suggesting that other factors reduce vectorial capacity at these locations.

3.4. Mating behavior and reproductive isolation between the molecular forms

The molecular forms mate assortatively as evident based on sperm identification in mated females from Mali (Tripet et al., 2001). Cross-mating (1.2%) was considerably lower than expected by chance (17%). Notably no post-zygotic isolation was detected in laboratory crosses of offspring of wild females (Diabate et al., 2007). Although the fitness of hybrids under natural conditions may be reduced, these results suggest that reproductive isolation between the molecular forms is primarily achieved by distinct mating behavior.

In many mosquitoes mating is initiated in swarms. Swarms are mostly seen at sunset and are composed of males flying 1–3 m above ground (often over a distinct marker), and forming a compact "cloud" less than one meter in diameter. Females fly into swarms and depart paired with a male (Charlwood and Jones, 1980; Charlwood et al., 2002; Diabate et al., 2006; Marchand, 1984). The mating behavior of *A. gambiae* is key to understand the mechanisms of reproductive isolation of sympatric populations of molecular forms and sibling species. The form composition in 26 swarms in Burkina Faso where the forms are sympatric revealed that males of both molecular forms were present in only four swarms (Diabate et al., 2006). This level of mixed swarms was considerably lower than that expected by chance based on the form composition indoors during the same time. These results suggest that spatial swarm segregation contributes to reproductive isolation, but that mate recognition within swarms is also needed to account for the low frequency of hybrids. In East Africa, Marchand (1984) reported mixed swarms between *A. gambiae* and *A. arabiensis* suggesting that mate recognition within swarms is widespread. Contrary to the findings from Burkina Faso, recent studies in Mali revealed a striking spatial segregation between sympatric molecular forms (Diabate et al. unpublished). In Mali, the S form swarms above bare ground, whereas the M form swarms above markers consisting of a dark center in a lighter background such as a well (Diabate et al. unpublished). The low mixing of swarms in Mali suggests that swarm segregation alone accounts for the assortative mating observed by Tripet et al. (2001). The possibility that different mechanisms facilitating assortative mating evolve in different populations (only 500 km apart) is intriguing (see below).

The role of flight tone in mate recognition has been evaluated between *A. gambiae* and *A. arabiensis* (Brogdon, 1998; Tripet et al., 2004; Wekesa et al., 1998) and between the molecular forms (Tripet et al., 2004). The overlap in flight tone of males and females between these species and forms was substantial, precluding it as the signal preventing cross-mating. Similar conclusion was reached in study of sibling species of *A. quadrimaculatus* (Caprio et al., 2001). Notably, all these studies measured flight tone of individuals which were not mate-seeking. In crane flies (Tipulidae), flight tone does not remain constant, but is changing as a "duet" between the male and the female that might lead to mating (Gibson and Russell, 2006). Analysis of flight tone in mate-seeking *A. gambiae* before copulation could determine if similar behaviors are exhibited by males and females of this species. Clearly, the challenge in measuring such individual 'songs' mixed in the chorus of the swarms may not be easy for the investigator or even the mosquitoes.

A preliminary study of cuticular hydrocarbons detected differences between chromosomal forms (Milligan et al., 1993), suggesting that olfaction is key to mate recognition. Weak but inconclusive evidence for differences in composition of cuticular hydrocarbons between the molecular forms were detected (Caputo et al., 2007). Additional evidence against olfactory cues involvement in mate recognitions was obtained from studies on indoor mating in Mali. Mark release recapture of virgin males and females in natural houses showed that mating occurred over a single day even when mosquitoes can leave the house through exit traps and without adaptation to laboratory conditions (Dao et al., 2008). Importantly, cross-mating between the molecular forms occurred indoors as much as mating between members of the same form (Dao et al., 2008), indicating that chemical cues such as pheromones and cuticular hydrocarbons do not play a major role in form recognition, unless such signals are only released or perceived during swarming. The results of this study suggest that indoors mating occurs naturally only in the M form.

4. Discussion

Recent advances in understanding divergence between the molecular forms and the mechanisms that facilitate assortative mating between them allow addressing several persisting questions about this speciation process and its consequences. These advances leave gaps in our knowledge and raise new questions awaiting future studies.

4.1. The ecological factors that promoted divergence

In most theoretical models of sympatric and parapatric speciation, divergent natural selection plays a dominant role (Genner et al., 2007; Grant and Grant, 1979; Schluter, 2001; Wu, 2001; Wu and Ting, 2004). Coluzzi et al. (2002) hypothesized that adaptation to different larval habitats was key to speciation in *A. gambiae* complex as exemplified by three independent speciation events that resulted in salty-water (brackish) tolerant species: *A. melas, A. merus*, and *A. bwambae*. Likewise, the association of the Bamako chromosomal form with rock pools along the tributaries of the Niger river strengthened that hypothesis (Manoukis et al., 2008; Toure et al., 1998) although specific ecological factors were not identified. Moreover, larval surveys in areas of sympatry failed to detect differences in habitat use by the molecular forms in Mali (Edillo et al., 2002) and by *A. gambiae* and *A. arabiensis* in Kenya (Gimnig et al., 2001). However, such surveys did not quantify the success of the forms in different habitats and did not estimate form and species relative abundance.

Consistent with that hypothesis, recent results suggest that the length of the hydroperiod of larval habitats is the principal ecological factor producing divergence between the molecular forms through selection mediated by predation and competition (Diabate et al., 2008). Field experiments demonstrated that the S form outperforms the M form in the absence of predators whereas the reverse is true if predator pressure increases. Predators are much more abundant in permanent larval habitats. These results explain the sharpest discontinuity exhibited by the molecular forms—the segregation between rice fields and surrounding savanna. Additionally, the affinity of the M form to aridity can also be explained by its ability to exploit permanent larval sites that sustain populations year round where rainfall is low. Many larval sites remain wet for several weeks and will harbor intermediate predator community (Williams, 2006), possibly allowing more equal success of both forms as suggested by high frequency of larval sites shared by both forms. That *A. arabiensis* in

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003

East Africa occupies (what appears to be) the "exact" niche filled by the M form in West Africa, suggests the same process operated in both cases, resulting in convergent evolution. However, *A. arabiensis* shows no population division (Kamau et al., 2007). The possible confounding effect of gene(s) conferring insecticide resistance on traits such as larval predator avoidance and longevity remains unknown because of the high frequency of the kdr resistance mutation in many populations of the S form, and because other genes conferring insecticide resistance, most of which are unknown, cannot be ruled out.

Desiccation and parasitism may have also mediated divergent selection because they are probably correlated with predation pressure as are other factors (Blaustein and Chase, 2007). Field evaluations of hybrid fitness in comparison to pure forms can be a powerful approach for this endeavor. Selection against hybrids is expected because their phenotype is intermediate and hence less adapted than one or the other pure form in each habitat. Ecological speciation against hybrids has been shown in a few divergent species such as threespine sticklebacks and butterfly *Heliconius erato* (Hatfield and Schulter, 1999; Mallet et al., 1998). In both cases, hybrids were viable in the laboratory, but showed lower fitness in natural conditions.

4.2. How did mate preference evolve?

Divergent natural selection acts on ecologically important traits rather than on mate selection. However, if these traits also affect mate choice, directly or indirectly, then reproductive isolation could evolve as a by-product of local adaptation (Rundle et al., 2000, 2005, Rice and Hostert, 1993, Dodd, 1989, Schluter, 2000, Nosil et al., 2002). Adaptation to permanent vs. temporary larval habitat, indirectly promotes spatial (and partly temporal) isolation between populations in arid and humid environments. Mutations accumulating in such isolated population may have modified mating behavior. Such mutations might be fixed by drift before populations reconnect, or once populations merged, reinforced by selection, because of reduced fitness of hybrids (Coyne and Orr, 1997; Genner et al., 2007; Hatfield and Schulter, 1999; Mallet et al., 1998).

This scenario may also explain the different mechanisms of assortative mating between populations in Mali (spatial swarm segregation) and Burkina Faso (within swarm recognition and spatial swarm segregation). Accordingly, isolation of diverging populations as described above occurred repeatedly and independent mechanisms facilitating assortative mating have evolved and are possibly spreading. It is also compatible with finding a third speciation island (on chromosome 2R) separating the molecular forms from Cameroon but not from Mali (Turner and Hahn, 2007). Such a process appears problematic because it could lead to additional subdivisions within form rather than a single division between forms. However, without reinforcement and in the face of gene flow, these mechanisms may be selected against rather quickly within form (between genotypes with adaptations to the same habitat) or spread rapidly within form due to their advantage in preventing cross-mating between forms.

4.3. Divergence riddles: which form was first? When they split? What is their future?

The answers to these questions remain highly speculative and should be considered cautiously. Few clues suggest that the S form preceded the M form, including its (1) larger geographical range (Fig. 1), (2) its location in the center of the continent where humid habitats abound and populations are more stable, and (3) within-form polymorphism across the 2R speciation islands in populations from Cameroon appears higher in the S form (Turner and Hahn, 2007), indicating that selection operated more extremely on the M form. Further, the scenario of parapatric speciation (above) suggests that isolated populations occurred in dryer areas where the M form predominates.

Genetic and phenotypic similarity indicates that divergence between the molecular forms is recent on evolutionary time scale. Coluzzi et al. (2002) hypothesized that the spread of agriculture led to the speciation of *A. gambiae* s.s. approximately 5000 years ago from a zoophilic ancestor, so the molecular forms must have split even later. Although plausible, there is little evidence to substantiate this hypothesis: (i) estimates of time based on molecular clocks cannot resolve dates on this scale (e.g., Krzywinski and Besansky (2003); (ii) molecular clock assumes no gene flow between the species, whereas available evidence suggests that gene flow between the forms is ongoing (Tripet et al., 2001; Turner and Hahn, 2007), and the length of time until gene flow between incipient species ceases is unknown (see below); (iii) the assumption that biting people, resting indoors, and growing in larval sites created by human activity required speciation rather than adaptation remains to be validated. Finally, variation in hydroperiod of natural larval sites existed regardless of man, hence its role in divergence does not support this hypothesis.

Not only their past, the future of the molecular forms is also unknown. Increasing divergence with time is most plausible, but since fusion and fission are possible and depend on ecological conditions, divergence may remain the same for long or even decrease. Such alternative scenarios were inferred for Darwin's finches and host races of plant-eating insects (Dres and Mallet, 2002; Grant and Grant, 1979).

4.4. Adaptive variation in A. gambiae

Positive selection was demonstrated by latitudinal and temporal clines in inversion frequencies (above) and from patterns of polymorphism and differentiation between the molecular forms operating on a "locus" on chromosome II (Turner and Hahn, 2007). Which trait was selected in that locus? What alternative states of the trait existed? And what ecological conditions promoted this 'event'? These questions, which are also relevant to polymorphic inversions of *A. gambiae*, highlight the enormous gap between genetic and ecological inferences in this species. The behavioral, physiological, and morphological adaptations that allow populations of *A. gambiae* to occupy diverse environments across Africa are mostly unknown. The molecular forms represent a promising opportunity to uncover adaptive variation, although local adaptation presumably exists within each form. Research on phenotypic traits under natural conditions (except on disease transmission and insecticide resistance) has been scanty, limiting interpretation of the wealth of genetic information that has been accumulating over recent years (e.g., Ribeiro et al., 2004; Stump et al., 2005; Turner et al., 2005). A central challenge to extend understanding of phenotypic variation is the measurement of fitness (or its key components), under natural conditions. As demonstrated by larval transplantation studies with and without predators (Diabate et al., 2005 vs. Diabate et al., 2008) and by the absence of mating specificity indoors (Dao et al., 2008) as opposed to outdoors (Tripet et al., 2001; Diabate et al., 2006 and unpublished), natural environment is often distorted even by semi-field conditions, resulting in misleading results. Screened 'greenhouses' set in natural conditions (e.g., Knols et al., 2002) are undoubtedly valuable but cannot replace natural studies.

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003

4.5. Public health implications

Available evidence suggests that the molecular (and chromosomal) forms are highly efficient malaria vectors, despite minor differences (above). These diverse vectors together increase the burden of malaria and lymphatic filariasis because they extend duration of disease transmission and its geographical range (Table 1 and Fig. 1). From disease transmission surveillance and control perspectives, the epidemiological importance of the subdivision of *A. gambiae* into molecular forms is modest or even low as argued by Curtis (2000). Because both contribute similarly to disease transmission, both are anthropophilic, biting at night, and indoor resting during the day—they can be targeted by similar means of adult control. Larval control methods might have a greater success in the fewer, larger, and predictable sites of the M form, but no such strategies are widely implemented. Of importance, however, is the need to monitor insecticide resistance in the two forms separately.

Indirectly, this compartmentalized vectorial system is far more robust given its higher capacity to adapt because of the high total effective population size and subdivision coupled with introgression (Mallet, 2005), leading to the spread of adaptive genes between "compartments". Future prospects for disease control is reduced not only because of the forms different susceptibility to insecticides but also because reproductive isolation will limit effectiveness of control strategies based on sterile male release or the spread of genetically engineered constructs reducing transmission (Benedict and Robinson, 2003), unless these populations are targeted simultaneously.

Acknowledgments

We thank Drs. Frederic Simard, Nicholas Manoukis, Cecilia Coscaron, Frederic Tripet, Adama Dao, Jen Hume, and Bill Hawley, for helpful comments and discussions on earlier versions of this paper. Special thanks to Drs. Nora Besansky, Nicholas Manoukis, Frederic Simard, and to Mr. Adama Dao and colleagues at MRTC, Mali for sharing unpublished data. This research was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Allergy and Infectious Diseases.

References

Ayala, F.J., Coluzzi, M., 2005. Chromosome speciation: humans, Drosophila, and mosquitoes. Proc. Natl. Acad. Sci. U.S.A. 102 (Suppl. 1), 6535–6542.
Bayoh, M.N., Thomas, C.J., Lindsay, S.W., 2001. Mapping distributions of chromosomal forms of *Anopheles gambiae* in West Africa using climate data. Med. Vet. Entomol. 15, 267–274.
Benedict, M.Q., Robinson, A.S., 2003. The first releases of transgenic mosquitoes: an argument for the sterile insect technique. Trends Parasitol. 19, 349–355.
Berticat, C., Duron, O., Heyse, D., Raymond, M., 2004. Insecticide resistance genes confer a predation cost on mosquitoes, *Culex pipiens*. Genet. Res. 83, 189–196.
Besansky, N.J., Krzywinski, J., Lehmann, T., Simard, F., Kern, M., Mukabayire, O., Fontenille, D., Toure, Y., Sagnon, N., 2003. Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. Proc. Natl. Acad. Sci. U.S.A. 100, 10818–10823.
Besansky, N.J., Lehmann, T., Fahey, G.T., Fontenille, D., Braack, L.E., Hawley, W.A., Collins, F.H., 1997. Patterns of mitochondrial variation within and between African malaria vectors, *Anopheles gambiae* and *A. arabiensis*, suggest extensive gene flow. Genetics 147, 1817–1828.
Black, W.C., Lanzaro, G.C., 2001. Distribution of genetic variation among chromosomal forms of *Anopheles gambiae* s.s: introgressive hybridization, adaptive inversions, or recent reproductive isolation? Insect Mol. Biol. 10, 3–7.
Blaustein, L., Chase, J.M., 2007. Interactions between mosquito larvae and species that share the same trophic level. Annu. Rev. Entomol. 52, 489–507.
Blaustein, L., Kiflawi, M., Eitam, A., Mangel, M., Cohen, J.E., 2004. Oviposition habitat selection in response to risk of predation in temporary pools: mode of detection and consistency across experimental venue. Oecologia 138, 300–305.
Brogdon, W.G., 1998. Measurement of flight tone differentiates among members of the *Anopheles gambiae* species complex (Diptera: Culicidae). J. Med. Entomol. 35, 681–684.
Bryan, J.H., Di Deco, M.A., Petrarca, V., 1982. Inversion polymorphism and incipient speciation in *Anopheles gambiae* s.s in the Gambia, West Africa. Genetica 59, 167–176.
Calzetta, M., Santolamazza, F., Carrara, G.C., Cani, P.J., Fortes, F., Di Deco, M.A., Torre, A.D., Petrarca, V., 2008. Distribution and chromosomal characterization of the *Anopheles gambiae* complex in Angola. Am. J. Trop. Med. Hyg. 78, 169–175.
Caprio, M.A., Huang, J.X., Faver, M.K., Moore, A., 2001. Characterization of male and female wingbeat frequencies in the *Anopheles quadrimaculatus* complex in Mississippi. J. Am. Mosq. Control Assoc. 17, 186–189.
Caputo, B., Dani, F.R., Horne, G.L., N'Fale, S., Diabate, A., Turillazzi, S., Coluzzi, M., Costantini, C., Priestman, A.A., Petrarca, V., della Torre, A., 2007. Comparative analysis of epicuticular lipid profiles of sympatric and allopatric field populations of *Anopheles gambiae* s.s. molecular forms and *A. arabiensis* from Burkina Faso (West Africa). Insect Biochem. Mol. Biol. 37, 389–398.
Chandre, F., Manguin, S., Brengues, C., Dossou Yovo, J., Darriet, F., Diabate, A., Carnevale, P., Guillet, P., 1999. Current distribution of a pyrethroid resistance gene (kdr) in *Anopheles gambiae* complex from west Africa and further evidence for reproductive isolation of the Mopti form. Parassitologia 41, 319–322.
Charlwood, J.D., Jones, M.D.R., 1980. Mating in the mosquito, *Anopheles gambiae*. s. l. II. Swarming behaviour. Physiol. Entomol. 5, 315–320.
Charlwood, J.D., Pinto, J., Sousa, C.A., Madsen, H., Ferreira, C., do Rosario, V.E., 2002. The swarming and mating behaviour of *Anopheles gambiae* s.s. (Diptera: Culicidae) from Sao Tome Island. J. Vector Ecol. 27, 178–183.
Cianchi, R., Villani, F., Toure, Y.T., Petrarca, V., Bullini, L., 1983. Electrophoretic study of different chromosomal forms within Anopheles gambiae s.s. Parassitologia 25, 239–241.
Coluzzi, M., Petrarca, V., Di Deco, M.A., 1985. Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae*. Bollettino Zool. 52, 45–63.
Coluzzi, M., Sabatini, A., 1967. Cytogenetic observations on species A and B of the *Anopheles gambiae* complex. Parassitologia 9, 73–87.
Coluzzi, M., Sabatini, A., della Torre, A., Di Deco, M.A., Petrarca, V., 2002. A polytene chromosome analysis of the *Anopheles gambiae* species complex. Science 298, 1415–1418.
Coluzzi, M., Sabatini, A., Petrarca, V., Di Deco, M.A., 1979. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. Trans. R. Soc. Trop. Med. Hyg. 73, 483–497.
Coyne, J.A., Orr, H.A., 1997. "Patterns of speciation in Drosophila" revisited. Evolution 51, 295–303.
Curtis, C.F., 2000. Infectious disease. The case for deemphasizing genomics in malaria control. Science 290, 1508.
Dao, A., Adamou, A., Yaro, A.S., Hamidou, M.M., Kassogue, Y., Traore, S., Lehmann, T., 2008. Assessment of alternative mating strategies in *Anopheles gambiae*: does mating occur indoors? J Med. Entomol. 45, 643–652.
Davidson, G., 1964. *Anopheles Gambiae*, a complex of species. Bull. World Health Organ. 31, 625–634.
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., 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., 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.
Di Deco, M.A., Petrarca, V., Villani, F., Coluzzi, M., 1980. Polimorphismo cromosomico da inversioni paracentriche ed eccesso degli eterocariotipi in ceppi di *Anopheles allevati* in laboaratorio. Parassitologia 22, 304–306.
Diabate, A., Baldet, T., Chandre, F., Akoobeto, M., Guiguemde, T.R., Darriet, F., Brengues, C., Guillet, P., Hemingway, J., Small, G.J., Hougard, J.M., 2002. The role of agricultural use of insecticides in resistance to pyrethroids in *Anopheles gambiae* s.l. in Burkina Faso. Am. J. Trop. Med. Hyg. 67, 617–622.
Diabate, A., Dabire, R.K., Heidenberger, K., Crawford, J., Lamp, W.O., Culler, L.E., Lehmann, T., 2008. Evidence for divergent selection between the molecular forms of *Anopheles gambiae*: role of predation. BMC Evol. Biol. 8, 5.
Diabate, A., Dabire, R.K., Kengne, P., Brengues, C., Baldet, T., Ouari, A., Simard, F., Lehmann, T., 2006. Mixed swarms of the molecular M and S forms of *Anopheles gambiae* (Diptera: Culicidae) in sympatric area from Burkina Faso. J. Med. Entomol. 43, 480–483.
Diabate, A., Dabire, R.K., Kim, E.H., Dalton, R., Millogo, N., Baldet, T., Simard, F., Gimnig, J.E., Hawley, W.A., Lehmann, T., 2005. Larval development of the molecular forms of *Anopheles gambiae* (Diptera: Culicidae) in different habitats: a transplantation experiment. J. Med. Entomol. 42, 548–553.
Diabate, A., Dabire, R.K., Millogo, N., Lehmann, T., 2007. Evaluating the effect of postmating isolation between molecular forms of *Anopheles gambiae* (Diptera: Culicidae). J. Med. Entomol. 44, 60–64.
Diuk-Wasser, M., Toure, M.B., Dolo, G., Bagayoko, M., Sogoba, N., Traore, S.F., Manoukis, N., Taylor, C.E., 2005. Vector abundance and malaria transmission in rice growing villages in Mali. Am. J. Trop. Med. Hyg. 72, 725–731.
Donnelly, M.J., Pinto, J., Girod, R., Besansky, N.J., Lehmann, T., 2004. Revisiting the role of introgression vs. shared ancestral polymorphisms as key processes shaping genetic diversity in the recently separated sibling species of the *Anopheles gambiae* complex. Heredity 92, 61–68.
Dres, M., Mallet, J., 2002. Host races in plant-feeding insects and their importance in sympatric speciation. Philos. Trans. R. Soc. Lond. B Biol. Sci. 357, 471–492.

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003

Edillo, F.E., Toure, Y.T., Lanzaro, G.C., Dolo, G., Taylor, C.E., 2002. Spatial and habitat distribution of *Anopheles gambiae* and Anopheles arabiensis (Diptera: Culicidae) in Banambani village, Mali. J. Med. Entomol. 39, 70–77.

Favia, G., della Torre, A., Bagayoko, M., Lanfrancotti, A., Sagnon, N., Toure, Y.T., Coluzzi, M., 1997. Molecular identification of sympatric chromosomal forms of *Anopheles gambiae* and further evidence of their reproductive isolation. Insect Mol. Biol. 6, 377–383.

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

Favia, G., Louis, C., 1999. Molecular identification of chromosomal forms of *Anopheles gambiae* sensu stricto. Parassitologia 41, 115–118.

Genner, M.J., Nichols, P., Carvalho, G.R., Robinson, R.L., Shaw, P.W., Smith, A., Turner, G.F., 2007. Evolution of a cichlid fish in a Lake Malawi satellite lake. Proc. Biol. Sci. 274, 2249–2257.

Gentile, G., Slotman, M., Ketmaier, V., Powell, J.R., Caccone, A., 2001. Attempts to molecularly distinguish cryptic taxa in *Anopheles gambiae* s.s. Insect Mol. Biol. 10, 25–32.

Gibson, G., Russell, I., 2006. Flying in tune: sexual recognition in mosquitoes. Curr. Biol. 16, 1311–1316.

Gimnig, J.E., Ombok, M., Kamau, L., Hawley, W.A., 2001. Characteristics of larval anopheline (Diptera: Culicidae) habitats in Western Kenya. J. Med. Entomol. 38, 282–288.

Grant, B.R., Grant, P.R., 1979. Darwin's finches: population variation and sympatric speciation. Proc. Natl. Acad. Sci. U.S.A. 76, 2359–2363.

Hatfield, T., Schulter, D., 1999. Ecological speciation in sticklebacks: environment-dependent hybrid fitness. Evolution 52, 866–873.

Hunt, R.H., Coetzee, M., Fettene, M., 1998. The *Anopheles gambiae* complex: a new species from Ethiopia. Trans. R. Soc. Trop. Med. Hyg. 92, 231–235.

Kamau, L., Munyekenye, G.O., Vulule, J.M., Lehmann, T., 2007. Evaluating genetic differentiation of Anopheles arabiensis in relation to larval habitats in Kenya. Infect. Genet. Evol. 7, 293–297.

Kesavaraju, B., Alto, B.W., Lounibos, L.P., Juliano, S.A., 2007a. Behavioural responses of larval container mosquitoes to a size-selective predator. Ecol. Entomol. 32 (June (3)), 262–272.

Kesavaraju, B., Damal, K., Juliano, S.A., 2007b. Threat-sensitive behavioral responses to concentrations of water-borne cues from predation. Ethology 113, 199–206.

Knols, B.G., Njiru, B.N., Mathenge, E.M., Mukabana, W.R., Beier, J.C., Killeen, G.F., 2002. MalariaSphere: a greenhouse-enclosed simulation of a natural *Anopheles gambiae* (Diptera: Culicidae) ecosystem in western Kenya. Malar. J. 1, 19.

Koella, J.C., Boete, C., 2002. A genetic correlation between age at pupation and melanization immune response of the yellow fever mosquito *Aedes aegypti*. Evolution 56, 1074–1079.

Krzywinski, J., Besansky, N.J., 2003. Molecular systematics of Anopheles: from subgenera to subpopulations. Annu. Rev. Entomol. 48, 111–139.

Lambrechts, L., Chavatte, J.M., Snounou, G., Koella, J.C., 2006. Environmental influence on the genetic basis of mosquito resistance to malaria parasites. Proc. Biol. Sci. 273, 1501–1506.

Lanzaro, G.C., Toure, Y.T., Carnahan, J., Zheng, L., Dolo, G., Traore, S., Petrarca, V., Vernick, K.D., Taylor, C.E., 1998. 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. 95, 14260–14265.

Lehmann, T., Dalton, R., Kim, E.H., Dahl, E., Diabate, A., Dabire, R., Dujardin, J.P., 2006. Genetic contribution to variation in larval development time, adult size, and longevity of starved adults of *Anopheles gambiae*. Infect. Genet. Evol. 6, 410–416.

Lehmann, T., Licht, M., Elissa, N., Maega, B.T., Chimumbwa, J.M., Watsenga, F.T., Wondji, C.S., Simard, F., Hawley, W.A., 2003. Population structure of *Anopheles gambiae* in Africa. J. Hered. 94, 133–147.

Mallet, J., 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20, 229–237.

Mallet, J., 2006. What does Drosophila genetics tell us about speciation? Trends Ecol. Evol. 21, 386–393.

Mallet, J., McMillan, W.O., Jiggins, C.D., 1998. Estimating the mating behavior of a pair of hybridizing Heliconius species in the wild. Evolution 52, 503–510.

Manoukis, N.C., Powell, J.R., Toure, M.B., Sacko, A., Edillo, F.E., Coulibaly, M.B., Traore, S.F., Taylor, C.E., Besansky, N.J., 2008. A test of the chromosomal theory of ecotypic speciation in *Anopheles gambiae*. Proc. Natl. Acad. Sci. U.S.A. 105, 2940–2945.

Manoukis, N.C., Toure, M.B., Sissoko, I., Doumbia, S., Traore, S.F., Diuk-Wasser, M.A., Taylor, C.E., 2006. Is vector body size the key to reduced malaria transmission in the irrigated region of Niono, Mali? J. Med. Entomol. 43, 820–827.

Marchand, R.P., 1984. Field observations on swarming and mating in *Anopheles gambiae* mosquitoes in Tanzania. Neth. J. Zool. 34, 367–387.

Masendu, H.T., Hunt, R.H., Govere, J., Brooke, B.D., Awolola, T.S., Coetzee, M., 2004. 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. 98, 393–396.

McKinnon, J.S., Mori, S., Blackman, B.K., David, L., Kingsley, D.M., Jamieson, L., Chou, J., Schluter, D., 2004. Evidence for ecology's role in speciation. Nature 429, 294–298.

Milligan, P.J., Phillips, A., Broomfield, G., Molyneux, D.H., Toure, Y.T., Coluzzi, M., 1993. A study of the use of gas chromatography of cuticular hydrocarbons for identifying members of the *Anopheles gambiae* (Diptera: Culicidae) complex. B. Entomol. Res. 83, 613–624.

Mukabayire, O., Caridi, J., Wang, X., Toure, Y.T., Coluzzi, M., Besansky, N.J., 2001. Patterns of DNA sequence variation in chromosomally recognized taxa of *Anopheles gambiae*: evidence from rDNA and single-copy loci. Insect Mol. Biol. 10, 33–46.

Nosil, P., Crespi, B.J., 2004. Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and sexual isolation in Timema cristinae walking-sticks. Evolution Int J Org Evolution 58, 102–112.

Nosil, P., Crespi, B.J., 2006a. Ecological divergence promotes the evolution of cryptic reproductive isolation. Proc. Biol. Sci. 273, 991–997.

Nosil, P., Crespi, B.J., 2006b. Experimental evidence that predation promotes divergence in adaptive radiation. Proc. Natl. Acad. Sci. U.S.A. 103, 9090–9095.

Persiani, A., Di Deco, M.A., Petrangeli, G., 1986. Osservazioni di laboratorio su polimorfismi da inversione originata da incrocitra popolazioni diverse de *Anopheles gambiae* s.s. Ann. Istit. Super Sanita 22, 221–224.

Ribbands, C.R., 1944. Differences between Anopheles melas and *Anopheles gambiae*. II. Salinity relations of larvae and maxillary palp banding of adult females. Ann. Trop. Med. Parasitol. 38, 87.

Ribeiro, J.M., Topalis, P., Louis, C., 2004. AnoXcel: an *Anopheles gambiae* protein database. Insect Mol. Biol. 13, 449–457.

Robert, V., Ouari, B., Ouedraogo, V., Carnevale, P., 1988. An ecologic study of adult and larval Culicidae in a rice field of Kou Valley, Burkino Faso. Acta Trop. 45, 351–359.

Rundle, H.D., Nagel, L., Wenrick Boughman, J., Schluter, D., 2000. Natural selection and parallel speciation in sympatric sticklebacks. Science 287, 306–308.

Rundle, H.D., Chenoweth, S.F., Doughty, P., Blows, M.W., 2005. Divergent selection and the evolution of signal traits and mating preferences. PLoS Biol 3, e368.

Schluter, D., 2001. Ecology and the origin of species. Trends Ecol. Evol. 16, 372–380.

Schwartz, A., Koella, J.C., 2002. Melanization of *Plasmodium falciparum* and C-25 sephadex beads by field-caught *Anopheles gambiae* (Diptera: Culicidae) from southern Tanzania. J. Med. Entomol. 39, 84–88.

Slotman, M.A., Mendez, M.M., Torre, A.D., Dolo, G., Toure, Y.T., Caccone, A., 2006. Genetic differentiation between the BAMAKO and SAVANNA chromosomal forms of *Anopheles gambiae* as indicated by amplified fragment length polymorphism analysis. Am. J. Trop. Med. Hyg. 74, 641–648.

Slotman, M.A., Tripet, F., Cornel, A.J., Meneses, C.R., Lee, Y., Reimer, L.J., Thiemann, T.C., Fondjo, E., Fofana, A., Traore, S.F., Lanzaro, G.C., 2007. Evidence for subdivision within the M molecular form of *Anopheles gambiae*. Mol. Ecol. 16, 639–649.

Stump, A.D., Fitzpatrick, M.C., Lobo, N.F., Traore, S., Sagnon, N., Costantini, C., Collins, F.H., Besansky, N.J., 2005. Centromere-proximal differentiation and speciation in *Anopheles gambiae*. Proc. Natl. Acad. Sci. U.S.A. 102, 15930–15935.

Toure, Y.T., Petrarca, V., Traore, S.F., Coulibaly, A., Maiga, H.M., Sankare, O., Sow, M., Di Deco, M.A., Coluzzi, M., 1994. Ecological genetic studies in the chromosomal form Mopti of *Anopheles gambiae* s.str. in Mali, west Africa. Genetica 94, 213–223.

Toure, Y.T., Petrarca, V., Traore, S.F., Coulibaly, A., Maiga, H.M., Sankare, O., Sow, M., Di Deco, M.A., Coluzzi, M., 1998. The distribution and inversion polymorphism of chromosomally recognized taxa of the *Anopheles gambiae* complex in Mali, West Africa. Parassitologia 40, 477–511.

Tripet, F., Dolo, G., Traore, S., Lanzaro, G.C., 2004. The "wingbeat hypothesis" of reproductive isolation between members of the *Anopheles gambiae* complex (Diptera: Culicidae) does not fly. J. Med. Entomol. 41 (May (3)).

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

Tripet, F., Wright, J., Cornel, A., Fofana, A., McAbee, R., Meneses, C., Reimer, L., Slotman, M., Thiemann, T., Dolo, G., Traore, S., Lanzaro, G., 2007. 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. 76, 81–87.

Turner, T.L., Hahn, M.W., 2007. Locus- and population-specific selection and differentiation between incipient species of *Anopheles gambiae*. Mol. Biol. Evol. 24, 2132–2138.

Turner, T.L., Hahn, M.W., Nuzhdin, S.V., 2005. Genomic islands of speciation in *Anopheles gambiae*. PLoS Biol. 3, e285.

Wang, R., Zheng, L., Toure, Y.T., Dandekar, T., Kafatos, F.C., 2001. 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. 98, 10769–10774.

Wekesa, J.W., Brogdon, W.G., Hawley, W.A., Besansky, N.J., C.G.A. Centers for Disease Control, Prevention, 1998. Flight tone of field-collected populations of *Anopheles gambiae* and *An. arabiensis* (Diptera: Culicidae). Physiol. Entomol. 23, 289–294.

Williams, D.D., 2006. The Biology of Temporary Waters. Oxford University Press, New York.

Wondji, C., Frederic, S., Petrarca, V., Etang, J., Santolamazza, F., Della Torre, A., Fontenille, D., 2005a. 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. 42, 998–1005.

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

Wondji, C., Simard, F., Lehmann, T., Fondjo, E., Same-Ekobo, A., Fontenille, D., 2005b. Impact of insecticide-treated bed nets implementation on the genetic structure

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003

of *Anopheles arabiensis* in an area of irrigated rice fields in the Sahelian region of Cameroon. Mol. Ecol. 14, 3683–3693.

Wu, C.-I., 2001. The genic view of the process of speciation. J. Evol. Biol. 14, 851–865.

Wu, C.I., Ting, C.T., 2004. Genes and speciation. Nat. Rev. Genet. 5, 114–122.

Yaro, A.S., Dao, A., Adamou, A., Crawford, J.E., Ribeiro, J.M., Gwadz, R., Traore, S.F., Lehmann, T., 2006a. The distribution of hatching time in *Anopheles gambiae*. Malar. J. 5, 19.

Yaro, A.S., Dao, A., Adamou, A., Crawford, J.E., Traore, S.F., Toure, A.M., Gwadz, R., Lehmann, T., 2006b. Reproductive output of female *Anopheles gambiae* (Diptera: Culicidae): comparison of molecular forms. J. Med. Entomol. 43, 833–839.

Yawson, A.E., McCall, P.J., Wilson, M.D., Donnelly, M.J., 2004. Species abundance and insecticide resistance of *Anopheles gambiae* in selected areas of Ghana and Burkina Faso. Med. Vet. Entomol. 18, 372–377.

Please cite this article in press as: Lehmann, T., Diabate, A., The molecular forms of *Anopheles gambiae*: A phenotypic perspective, Infect. Genet. Evol. (2008), doi:10.1016/j.meegid.2008.06.003