(*27*). In the beetle *Tribolium castaneum*, the three genes have a similar expression pattern but *pb* expression is not diminished in a *Dfd* mutant, suggesting that *Dfd* is not necessary to activate *pb* in this species (*28*). This probably reflects a difference in *pb* regulatory sequences, as the *Tribolium* Deformed protein can activate *Drosophila pb* when it is expressed in *Drosophila* (*29*). In the bug *Oncopeltus fasciatus*, *Dfd* cannot activate *pb* because *pb* is not even present in the maxillae (*30*). Additionally, although *pb* and *Scr* are coexpressed in the labial appendages, RNA interference analysis suggests that *Scr* does not activate *pb* (*31*). Lastly, even in the *Drosophila* adult the regulatory hierarchy appears to be different from that in the *Drosophila* embryo; *Scr* does not activate *pb* in adults, but rather *pb* is necessary to activate *Scr* (*32*). Thus, we have three insect species and four different regulatory systems to control the expression of *proboscipedia*. Considering the millions of different insect species, these results suggest enormous diversity in the regulation of this, and other, developmental genes.

The completed sequences of the *Drosophila* and *Anopheles* genomes and the prospective sequencing of the *Apis* and *Aedes* genomes will provide significant insights into the insects and their development, behavior, and evolution (Fig. 1). But these four species represent only the beginning of an analysis of the Insecta, much less of the whole of the Arthropoda. Next we might consider sequencing genomes of representatives from the Coleoptera and Lepidoptera. These two insect orders contain many of our most serious agricultural pests and, together with the Diptera and Hymenopterá, comprise the "Big Four" insect orders that have evolved "complete" (holometabolous) development. The scientific community is now blessed with a wealth of sequencing capacity. Given the obvious importance of insects to our well-being and existence, it is important that some of it be used to build a strong empirical foundation for comparative insect genomics.

References

1. R. A. Holt *et al.*, *Science* **298**, 129 (2002).
2. M. D. Adams *et al.*, *Science* **287**, 2185 (2000).
3. S. L. Hoffman, G. M. Subramanian, F. H. Collins, J. C. Venter, *Nature* **415**, 702 (2002).
4. World Health Organization, *The World Health Report 2000, Health Systems: Improving Performance* (World Health Organization, Geneva, 2000).
5. B. J. Beaty, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10295 (2000).
6. K. S. Aultman *et al.*, *Trends Parasitol.* **17**, 507 (2001).
7. F. Catteruccia *et al.*, *Nature* **405**, 959 (2000).
8. M. L. Allen *et al.*, *J. Med. Entomol.* **38**, 701 (2001).
9. Z. N. Adelman *et al.*, *Mol. Biochem. Parasitol.* **121**, 1 (2002).
10. J. Ito *et al.*, *Nature* **417**, 452 (2002).
11. D. W. Severson *et al.*, *Annu. Rev. Entomol.* **46**, 183 (2001).
12. K. S. Rai *et al.*, *Adv. Genet.* **41**, 1 (1999).
13. I. V. Sharakhov *et al.*, *Science* **298**, 182 (2002).
14. E. M. Zdobnov *et al.*, *Science* **298**, 149 (2002).
15. J. R. Anderson *et al.*, *Mol. Phylogenet. Evol.* **20**, 316 (2001).
16. D. W. Severson, personal communication.
17. E. Pennisi, *Science* **296**, 1589 (2002).
18. G. E. Robinson, *Trends Ecol. Evol.* **14**, 202 (1999).
19. ———, *Science* **297**, 204 (2002).
20. B. P. Berman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 757 (2002).
21. G. G. Loots *et al.*, *Science* **288**, 136 (2000).
22. K. Ishihara *et al.*, *Genome Res.* **10**, 664 (2000).
23. D. A. Papatsenko *et al.*, *Genome Res.* **12**, 470 (2002).
24. R. C Brusca, G. J. Brusca, *Invertebrates* (Sinauer Associates, Sunderland, MA, 1990).
25. W. C. Wheeler *et al.*, *Cladistics* **17**, 113 (2001).
26. E. Abouheif, G. A. Wray, *Science* **297**, 249 (2002).
27. D. B. Rusch, T. C. Kaufman, *Genetics* **156**, 183 (2000).
28. M. DeCamillis *et al.*, *Genetics* **159**, 1643 (2001).
29. S. J. Brown *et al.*, *Dev. Genes Evol.* **209**, 389 (1999).
30. B. T. Rogers *et al.*, *Evol. Dev.* **4**, 1 (2002).
31. C. L. Hughes *et al.*, *Development* **127**, 3683 (2000).
32. A. Abzhanov *et al.*, *Development* **128**, 2803 (2001).

VIEWPOINT

Speciation Within *Anopheles gambiae*—the Glass Is Half Full

A. della Torre,[1]* C. Costantini,[1]* N. J. Besansky,[3] A. Caccone,[4] V. Petrarca,[2] J. R. Powell,[4] M. Coluzzi[1]

Restrictions to gene flow among molecular forms of the mosquito *Anopheles gambiae sensu stricto* reveal an ongoing speciation process affecting the epidemiology of malaria in sub-Saharan Africa.

The most important vector of the malaria parasite in sub-Saharan Africa is the mosquito *Anopheles gambiae sensu stricto* (*s.s.*). It belongs to a group of sibling species—known as the *A. gambiae* complex—that are morphologically indistinguishable but exhibit distinct genetic and eco-ethological differences reflected in their ability to transmit malaria. *Anopheles gambiae s.s.* shows extreme genetic heterogeneity, revealed not only by the traditional study of chromosomal inversions (*1*) but also by recent studies of molecular markers such as X-linked ribosomal DNA (rDNA). So far, extensive molecular analyses have attempted to distinguish the number of isolated or semi-isolated genetic units of *A. gambiae s.s.* that exist and whether these are evolving into separate species (speciation). Elucidating the genetic population structure of the *A. gambiae s.s.* complex is a prerequisite for determining which genetic units of the complex are the vectors of malaria, and unraveling the ecological and ethological differences that are relevant to disease transmission. Such knowledge will improve our understanding of malaria epidemiology and will help in implementing appropriate vector control strategies.

Genotyping X-linked rDNA of *A. gambiae s.s.* has led to the characterization of two molecular forms (M and S) that differ in both the transcribed and nontranscribed spacers in the rDNA repeat unit (*2*–*4*). The relationship between the M and S molecular forms and the chromosomal forms—defined according to nonrandom associations of inversions in chromosome 2 (*1*)—varies according to their ecological and geographic distribution (Fig. 1). In some areas of West Africa (for example, Mali and Burkina Faso), there is a one-to-one correspondence between the M molecular form and the Mopti chromosomal form. Similarly, the S molecular form always corresponds to the Savanna or Bamako chromosomal form (*5*). In other areas of West Africa, this clear correspondence breaks down (*2*). For example, in populations inhabiting forests or humid savannas, both molecular forms are characterized by high frequencies of the standard arrangement in chromosome 2 indicative of the Forest chromosomal form. Within the S form, a small proportion show ambiguous cytological configurations, indicating the presence of chromosome 2 arrangements typical of chromosomal forms other than Savanna and Bamako. Outside Mali and Burkina Faso, the M form may exhibit chromosomal arrangements typical of the Bissau, Savanna, or Forest forms.

[1]Parasitology Unit, Department of Public Health, [2]Department of Genetics and Molecular Biology, University of Rome "La Sapienza," P.le Aldo Moro 5, 00185 Rome, Italy. [3]Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA. [4]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA.

*To whom correspondence should be addressed. E-mail: ale.dellatorre@uniroma1.it; carlo.costantini@uniroma1.it

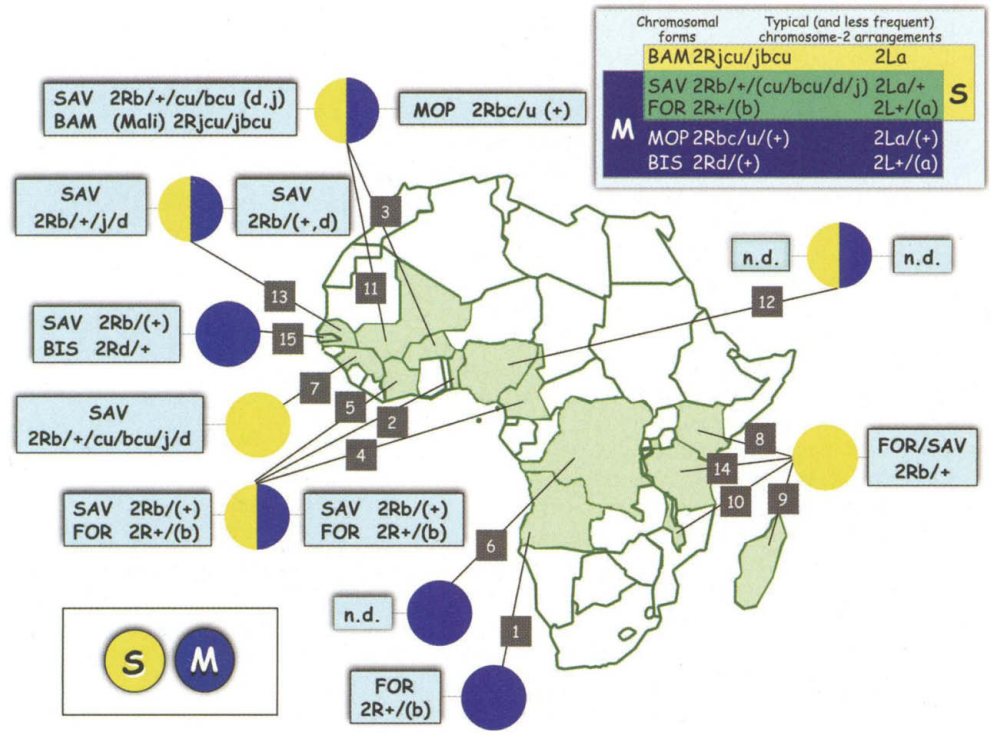

Fig. 1. Geographical distribution of the molecular forms (S and M) of *Anopheles gambiae sensu stricto* and their relation to chromosomal forms. Chromosomal inversions characterizing each chromosomal form (*1*) are listed first, followed by less frequent arrangements listed in parentheses. Two-color circles indicate the presence of one or both molecular forms (*2, 4, 9, 31, 32*) without reference to relative frequencies; absence of one form refers only to lack of detection in the mosquito samples analyzed, not to its actual absence in the area. Chromosomal forms: FOR, Forest; SAV, Savanna; MOP, Mopti; BAM, Bamako; BIS, Bissau; n.d., karyotype not determined. African countries: 1, Angola; 2, Benin; 3, Burkina Faso; 4, Cameroon; 5, Côte d'Ivoire; 6, Democratic Republic of the Congo; 7, Guinea; 8, Kenya; 9, Madagascar; 10, Malawi; 11, Mali; 12, Nigeria; 13, Senegal; 14, Tanzania; and 15, the Gambia.

The importance of chromosomal inversions in ecological adaptation has been well established [reviewed in (*6*)], suggesting that the different chromosomal forms are indicators of adaptation to different ecological habitats. In contrast, the distinct M and S molecular forms reflect barriers to gene flow indicative of incipient speciation (*2*). Extensive analyses of DNA regions other than rDNA initially failed to show consistent sequence differences corresponding to the M and S molecular forms (*4, 7*). However, recent DNA-based data are emerging in support of the M and S distinctions. For example, although most microsatellite loci show similar allelic frequencies in M and S forms, differences in allelic frequencies indicate restricted gene flow in both Mali (*8*) and Cameroon (*9*). Two microsatellite loci near the centromeric region of the X chromosome exhibit very distinct differences between M and S forms in Mali (*10*). The *kdr* allele in the *para* sodium channel gene, which confers resistance to pyrethroid insecticides, is found in S form populations from several West African countries; it could not be detected in M form populations from the same locales (*2, 11, 12*), with the single exception of Benin (*13*). Sequence analysis of intron I upstream of the *kdr* mutation has shown that S and M populations across West Africa are consistently different at one nucleotide (*14–16*). Furthermore, rDNA analysis in the closely related sibling species *A. arabiensis* would argue against the hypothesis that rDNA differences between the molecular forms of *A. gambiae* are due to unusual evolutionary dynamics (for example, very rapid concerted evolution) (*17*).

Although interbreeding between M and S forms yields fertile progeny, M-S hybrids are rarely observed in nature. Where these forms overlap in time and space, the rate of heterogamous insemination is ~1% (*18*), clearly demonstrating the existence of a premating barrier, albeit an incomplete one. Thus, both indirect and direct genetic evidence indicates incomplete but substantial barriers to gene flow between different *A. gambiae s.s.* molecular forms. Does ongoing gene flow signify a glass half empty, or does a premating barrier offer a glass half full? The data suggest that we are observing speciation at its very earliest stages, with the persistence of variation shared because of recent common ancestry and with low levels of gene flow continuing to homogenize regions of the genome not directly involved in the speciation process. This may explain why a random selection of nucleotide sequences reveals no differentiation, in contrast to the recently emerged *kdr* allele and the more rapidly evolving microsatellite markers and rDNA. In an attempt to reconcile differences among data sets, Gentile *et al.* (*17*) proposed that S and M forms "may have mosaic genomes consisting of parts completely differentiated between which gene flow is barred, whereas other parts of the genome are free to pass between forms." If correct, this suggests that a debate over taxonomic status hinging on absolute levels of gene flow will lack biological relevance and distract from the main issues: how restrictions on gene flow affect the ecology and behavior of the molecular forms; to what extent and under what circumstances they hinder the circulation of traits such as insecticide resistance or the possible spread of transgenes; and whether genetically engineered mosquitoes should be used in vector control measures (*19, 20*).

The vectorial potency of *A. gambiae s.s.* stems from its strong association with humans, that is, its preference for biting humans exacerbated by its capacity to exploit changes in its natural habitat induced by *Homo sapiens*. There is evidence (*21, 22*) that in marginal arid environments of Burkina Faso—where the Savanna and Mopti chromosomal forms correspond to the S and M molecular forms, respectively— these two taxa contrast significantly in the way they exploit limiting resources, such as larval breeding and adult resting habitats. Ecoethological differences regarding their degree of association with the human domestic environment when biting and resting are under investigation. The M molecular form shows the closest association with the domestic environment and larval habitats created by human activities, whereas the S form is more frequent in rain-dependent temporary breeding sites (*21, 22*). This confirms what was inferred from differential microgeographic distributions for the Mopti and Savanna chromosomal forms (*1, 23, 24*).

These observations provide us with further clues to the nature and mechanisms of the speciation process. The occupation by the M form of relatively recent ecological niches produced by human-made modifications of the environment in marginal habitats has created new opportunities for spe-

cialization and the avoidance of intraspecific competition. This selective force is presumably driving the speciation process. It has been proposed that co-adapted chromosomal inversions are crucial for establishing populations in marginal habitats that could lead to the formation of new species, although the inversions per se are not the cause of the evolution of subsequent barriers to gene flow (*25*).

The taxonomic and genetic complexity of *A. gambiae s.s.* has serious consequences for malaria transmission. The ongoing speciation process leading to the M form has extended the transmission potential of this vector in space and time (*23, 24*). In dry areas of West Africa where malaria is hyper- to holoendemic (*26*), this taxon is able to exploit breeding opportunities due to human activities that would otherwise be available only to *A. arabiensis*; such is the case in areas of Eastern Africa with a similar climate (like northern Sudan) where *A. gambiae s.s.* is absent and malaria is hypo- to mesoendemic (*27*). Moreover, in dry savannas, the ability of the M form to breed year-round in permanent human-dependent larval habitats extends the malaria transmission period well beyond the rainy season, when the S form apparently disappears (*28*). Analogous situations are seen with other Afrotropical malaria mosquito vectors such as *A. funestus*, which has two West African chromosomal forms (Folonzo and Kiribina) that clearly differ in their degree of contact with humans and therefore have quite different vectorial potentials (*29*). It is likely that in both *A. gambiae* and *A. funestus*, chromosomal inversions allow more specialized and therefore more efficient exploitation of both spatial and temporal environmental heterogeneity. This is expected to have implications for such traits as the survival probability of individual mosquitoes and the stability of vector populations, both important features of malaria epidemiology (*30*).

The complete genome sequence of *A. gambiae* will deepen our understanding of the process of adaptation and speciation of this insect vector. One immediate application of this information, already in progress, is the cloning of inversion breakpoints on chromosome 2. Comparative analysis of the sequences across and surrounding each breakpoint will allow us to identify and study the gene clusters protected by recombination and may yield clues about the origin of inversions and their importance.

The concentration within four closely related species of the *A. gambiae* complex (*A. gambiae*, *A. arabiensis*, *A. melas*, and *A. merus*) of several inversions along the central and subtelomeric sections of the 2R chromosomal arm is unlikely to be coincidental. These inversions may be associated with genome regions that encode traits of ecological and behavioral importance. The availability of the entire *A. gambiae* genome will facilitate polymerase chain reaction–based assays that will complement laborious karyotyping of semigravid adult females, providing new opportunities for field studies on mosquito ecology and behavior. A long-term goal is gene discovery using a complete genome chip. The very recent divergence of the *A. gambiae s.s.* molecular forms and the likelihood that only a few genes are involved in reproductive isolation and ecological diversification means that the entire *A. gambiae* genome will have to be screened in order to identify differences in gene sequence and coordinated gene expression between incipient species.

A. gambiae provides us with an exceptional opportunity to observe evolution in action, potentially operating over the time frame of the thousands of years since humans began to modify the Afrotropical ecosystem (*1, 6, 24*). The buildup of barriers to gene flow during the speciation process resulting in separation of the molecular forms of *A. gambiae* can be compared to a glass half full. Now, we must fully elucidate the mechanisms and dynamics of evolutionary change in *A. gambiae* populations—information that will be essential if we are ever to control this nefarious insect vector.

References and Notes

1. M. Coluzzi *et al.*, *Science*, in press; published online 3 October 2002 (10.1126/science.1077769).
2. A. della Torre *et al.*, *Insect Mol. Biol.* **10**, 9 (2001).
3. G. Favia *et al.*, *Insect Mol. Biol.* **10**, 19 (2001).
4. G. Gentile *et al.*, *Insect Mol. Biol.* **10**, 25 (2001).
5. G. Favia *et al.*, *Insect Mol. Biol.* **6**, 377 (1997).
6. J. R. Powell *et al.*, *Parassitologia* **41**, 101 (1999).
7. O. Mukabayire *et al.*, *Insect Mol. Biol.* **10**, 33 (2001).
8. C. Taylor *et al.*, *Genetics* **157**, 743 (2001).
9. C. Wondji *et al.*, *Insect Mol. Biol.* **11**, 11 (2002).
10. R. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10769 (2001).
11. F. Chandre *et al.*, *Parassitologia* **41**, 319 (1999).
12. C. Fanello *et al.*, in preparation.
13. C. Fanello *et al.*, *Trans. R. Soc. Trop. Med. Hyg.* **94**, 132 (2000).
14. M. Weill *et al.*, *Insect Mol. Biol.* **9**, 451 (2000).
15. F. Santolamazza *et al.*, *Parassitologia* **44**, 165 (2002).
16. In Benin, both the *kdr* allele and the "S-type" intron I were acquired by M from S.
17. G. Gentile *et al.*, *Genetics* **161**, 1561 (2002).
18. F. Tripet *et al.*, *Mol. Ecol.* **10**, 1725 (2001).
19. T. W. Scott, W. Takken, B. G. J. Knols, C. Boëte, *Science* **298**, 117 (2002).
20. L. Alphey *et al.*, *Science* **298**, 119 (2002).
21. N. F. Sagnon *et al.*, *Parassitologia* **42**, 157 (2000).
22. C. Costantini *et al.*, unpublished data.
23. Y. T. Touré *et al.*, *Genetica* **94**, 213 (1994).
24. Y. T. Touré *et al.*, *Parassitologia* **40**, 477 (1998).
25. M. Coluzzi, in *Mechanisms of Speciation*, C. Barigozzi, Ed. (Liss, New York, 1982), pp. 143–153.
26. The degree of malaria endemicity in an area is classified according to the proportion of children with a palpably enlarged spleen. Increasing levels of endemicity are expressed as hypoendemic (0 to 10%), mesoendemic (11 to 50%), hyperendemic (>50%), or holoendemic (>75%).
27. A. R. Zahar, *Vector Bionomics in the Epidemiology and Control of Malaria. Part I. The WHO African Region & the Southern WHO Eastern Mediterranean Region. Section III: Vector Bionomics, Malaria Epidemiology and Control by Geographical Areas. (D) East Africa; (E) Eastern Outer Islands; (F) Southwestern Arabia VBC/85.3 MAP/85.3* (World Health Organization, Geneva, 1985).
28. Y. T. Touré *et al.*, *Med. Vet. Entomol.* **10**, 197 (1996).
29. C. Costantini *et al.*, *Parassitologia* **41**, 595 (1999).
30. M. Coluzzi, *Bull. WHO* **62**, 107 (1984).
31. G. C. Carrara *et al.*, *Parassitologia* **44**, 43 (2002).
32. T. Lehmann, unpublished data.

VIEWPOINT

The Ecology of Genetically Modified Mosquitoes

Thomas W. Scott,[1]* Willem Takken,[2] Bart G. J. Knols,[2] Christophe Boëte[3]

Ecological and population biology issues constitute serious challenges to the application of genetically modified mosquitos (GMM) for disease control.

Optimism that mosquito-borne diseases such as malaria, dengue, and filariasis can be effectively controlled or even eradicated with inexpensive drugs, vaccines, or insecticides has been sorely tested (*1*). The impact of drugs is debatable, vaccine development is slow, and mosquitoes are becoming resistant to insecticides, including those used to treat bed nets (*2*). Such shortfalls have been used to justify research on mosquito population replacement—that is, the release into natural mosquito populations of genetically modified mosquitoes (GMM) rendered refractory to pathogen infection—to reduce or eliminate disease transmission (*3*).

A big hurdle to battling vector-borne diseases is our incomplete understanding of parasite transmission ecology, which is hampering GMM efforts in particular and public health initiatives in general. The GMM strategy should serve as a case study for ways to improve overall disease prevention, because the