# Evolutionary studies of malaria vectors

Martin J. Donnelly, Frédéric Simard and Tovi Lehmann

**The rationales given for studies of the population genetics of vectors are usually: (1) to predict the spread of genes, such as genes conferring insecticide resistance or possibly refractoriness to parasites and (2) to reveal novel insights into the epidemiology and transmission of vector-borne disease. The successful genetic transformation of mosquitoes has highlighted the need for a critical assessment of the rapidly accumulating body of data on the population genetics of malaria vectors. This article assesses how successful molecular genetic techniques have been in revealing new population patterns.**

The population structure of malaria vectors, like that of other organisms, reflects the combined actions of contemporary GENE FLOW (see Glossary), GENETIC DRIFT, selection and demographic history. Resolving the relative importance of each of these influences and extracting the unique effect of gene flow has been the theme for most population genetics work in the past decade. However, do ecological similarities between species allow the inference of a population structure that can be regarded as typical for primary vectors? Is it possible to determine what are the major architects of population structure in these systems? These questions and the problems of estimating contemporary gene flow in *Anopheles* vectors will be discussed.

## Population structure of primary malaria vectors

The most thoroughly studied anopheline species are the primary vectors of malaria. Within this group, the *Anopheles gambiae* complex has received the most attention. Primary vectors, similar to other pest species, have several common ecological attributes including: (1) wide geographical distribution; (2) high local abundance, although often seasonal; (3) good dispersal and colonizing ability; (4) adaptations to exploit different man-made environments (e.g. buildings, rice fields and vehicles); (5) a strong preference to feed on human blood and (6) a high susceptibility to human pathogens. Before the agricultural revolution, ~5000–10 000 years ago, human densities and resources were probably too scarce to result in such specialization; therefore, it is believed that most primary vectors have attained pest status only recently. The genetic structure of populations is shaped by dispersal (the key determinant of gene flow) and abundance (the key determinant of EFFECTIVE POPULATION SIZE). Therefore, the ecological similarities suggest that the typical population structure of primary vectors is probably shallow, with a weak effect of distance on differentiation. Homogeneity across vector populations is expected based on: (1) moderate to high gene flow inferred from the high mobility of these insects, (2) large effective population size and (3) the common diversity, which all current populations inherited from the recently expanded original stock.

Empirical studies do indeed show evidence for a shallow population structure and, in particular, the weak effect of distance on differentiation. These structuring patterns were observed in all three principal African vectors: *An. gambiae s.s.* [1], *Anopheles arabiensis* [2] and *Anopheles funestus* [3], in the Asian *Anopheles dirus* A and D [4], and *Anopheles maculatus* [5], and in the Southern and Central American vectors: *Anopheles darlingi* [6], *Anopheles pseudopunctipennis* [7] and *Anopheles albimanus* [8] (Table 1). A deep

**Martin J. Donnelly***
Division of Parasite and Vector Biology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK L3 5QA.
*e-mail: mjames@liv.ac.uk

**Frédéric Simard**
Laboratoire de Lutte contre les Insectes Nuisibles (LIN), Institut de Recherche pour le Développement (IRD), 911, Avenue Agropolis – BP 5045, 34032 Montpellier Cedex 1, France.

**Tovi Lehmann**
Centers for Disease Control and Prevention, MS F22, 4770, Buford Hwy, Chamblee, GA 30341, USA.

### Glossary

**Allopatric:** Species having separate and mutually exclusive geographical distributions.
**Assortative mating:** Non-random mating within a population where individuals tend to mate with individuals resembling themselves (positive assortative mating) or with those that differ in certain character(s) (negative assortative mating).
**Deme:** A local population unit of a species within which mating is random.
**Effective population size ($N_e$):** A measure of genetic drift that can be expressed as the number of parents that contribute gametes to the next generation after adjustment for their reproductive success (i.e. sexually mature individuals that mate successfully and transmit their genes to the next generation). $N_e$ is defined as the size of an idealized population (hypothetically PANMICTIC, not subject to selection, with non-overlapping generations, in which the sex ratio is one, and each reproductive individual has the same expected reproductive success) that experiences genetic drift at the same rate as the natural population under study. In a series of generations, each with a different $N_e$, the overall estimate will approximate the lowest $N_e$ (as the harmonic mean).
**$F_{ST}$:** A measure of differentiation between populations based on the interpopulation component of total genetic variation
**Gene flow ($N_m$):** The spread of alleles or genes as a result of mating between individuals from different populations.
**Genetic drift:** Random changes in allele or gene frequencies in populations that occur over time as a result of the finite number of gametes from the parent generation that form the subsequent generation. This process does not involve selection.
**Inversions:** Chromosomal rearrangement (such as on the right arm of chromosome 2 in *Anopheles gambiae s.l.*) in which a segment of the chromosome was cut, flipped and pasted back. Hence, the sequence of genes appears in reverse of the original orientation.
**Panmictic:** A state of random mating resulting in a homogenous population.
**Sympatric:** Species occupying fully or partially overlapping habitats.

http://parasites.trends.com 1471-4922/02/$ – see front matter  PII: S1471-4922(01)02198-5

**Table 1. Comparison of estimates of differentiation across large geographical distances[a]**

| Species | Vector status | Geographic distance (km) | $F_{ST}$ or $R_{ST}$ value[b] | Marker system | Refs |
|---|---|---|---|---|---|
| **Africa** | | | | | |
| *Anopheles gambiae*[c] | 1° | 6000 | 0.034 | STR | [1] |
| *An. gambiae*[c] | 1° | 6000 | 0.036 | Allo | [1,55] |
| *An. gambiae*[c] | 1° | 7000 | 0.085 | mtDNA | [11] |
| *Anopheles arabiensis*[c] | 1° | 7000 | 0.044 | mtDNA | [11] |
| *An. arabiensis*[c] | 1° | 3500 | 0.035 | STR | [2] |
| **Asia** | | | | | |
| *Anopheles dirus* A[c] | 1° | 1500 | 0.018 | mtDNA | [34] |
| *An. dirus* A[c] | 1° | 1500 | 0.02 | STR | [4] |
| *An. dirus* D[c] | 1° | 1500 | 0.022 | mtDNA | [34] |
| *An. dirus* D[c] | 1° | 1500 | 0.07 | STR | [4] |
| *An. dirus* C | 2° | 700 | 0.41 | mtDNA | [34] |
| *An. dirus* C | 2° | 700 | 0.19 | STR | [4] |
| *Anopheles maculatus* | 2° | 1100 | 0.051 | STR | [5] |
| **Europe** | | | | | |
| *Anopheles claviger* | | | | | |
| Group I | No[d] | 460 | 0.011 | Allo | [56] |
| Group II | No[d] | 430 | 0.011 | Allo | [56] |
| **South and Central America** | | | | | |
| *Anopheles darlingi* | 1° | 4000 | 0.10 or 0.64[e] | Allo or mtDNA | [6,57] |
| *Anopheles pseudopunctipennis* | 2° | 7000 | 0.38 | Allo | [7] |
| *Anopheles albimanus* | 1° | 2500 | 0.155 | mtDNA | [8] |
| **North America** | | | | | |
| *Anopheles punctipennis* | No[d] | 130 | 0.11 | mtDNA | [29] |
| **Non-anophelines** | | | | | |
| *Culex pipiens* | NA | 1000 | 0.036 | Allo | [13] |

[a]Estimates involving island populations or populations under intense vector-control were excluded (the low values of differentiation observed in primary vectors). Abbreviations: Allo, allozymes; mtDNA, mitochondrial DNA; NA, not applicable; STR, simple tandem repeat.
[b]$F_{ST}$ and $R_{ST}$ measure the differentiation between populations based on the interpopulation component of the total genetic variation [24]. Values above 0.2 are indicative of high levels of structuring.
[c]Species in which the multi-locus tests detected a recent population expansion signature.
[d]Species in areas where malaria is not endemic.
[e]An anomalously high value was recorded, possibly indicating a comparison between cryptic species.

subdivision of populations was only demonstrated for the Asian *An. dirus* C [4], and the Southern American *Anopheles nuneztovari* and *Anopheles aquasalis* (cited in Ref. [9]), all of which are regionally important (i.e. secondary) vectors. However, it should be noted that there is a need for more studies on secondary vectors to enable a more detailed comparison. Population structuring in primary vectors is possibly even lower than initial analyses suggest because the values of differentiation are often upwardly biased by the inclusion of a single uncharacteristic population. The atypical population is usually separated from other populations by a small geographic distance, and appears to be unique because it might have experienced a founder effect, extensive insecticide control or is geographically isolated. In *An. gambiae s.s*, for example, the highest differentiation was detected between two populations which were <700 km apart but separated by the Rift Valley Complex, whereas significantly lower differentiation has been measured between populations separated by 6000 km [1,10,11]. Similarly, in *An. albimanus*, the most distinct population was in Panama, in the centre of the study area, whereas more distant populations were relatively similar [8]. The high overall genetic differentiation within *An. pseudopunctipennis* was mainly a result of the inclusion of populations from the Caribbean island of Grenada and from northern Chile [7]. In each of these cases, the lower genetic diversity in the deviant populations suggests that they had experienced a population contraction or a bottleneck. Similar levels of population differentiation were also observed in non-anophelines, such as in worldwide populations of *Aedes aegypti* (see Ref. [12] and references cited therein) and *Culex pipiens* [13]. In both species, differentiation between populations was low (despite putative subdivision owing to large insecticide-spray campaigns) probably as a result of extensive gene flow mainly mediated by human transportation that replenished diversity from neighbouring uncontrolled regions.

Nonetheless, the low-population differentiation revealed by multi-locus microsatellite studies and sequence-based studies of mitochondrial DNA (mtDNA) typically reported in anophelines contrasts with studies using chromosomal markers, which suggest sharp discontinuities between populations that are often SYMPATRIC. There is strong evidence that the frequencies of certain inversion arrangements are shaped by selection [14,15]. The most notable case is *An. gambiae s.s.* where Coluzzi and colleagues [14,15] demonstrated that, in West Africa, several taxa of

uncertain status, principally defined by polymorphic INVERSIONS on the right arm of chromosome two (2R), exist in sympatry. Similar patterns of chromosomal forms were also found in *An. funestus*, but population subdivision was not supported by subsequent mtDNA and ribosomal DNA (rDNA) analysis [3,16–18]. In *An. darlingi*, fixed chromosomal differences have also been reported [19] but, despite documented heterogeneities in behaviour, physiology and ecology [20], combined analysis of molecular and morphological markers showed only weak differentiation between populations throughout the species range (but see Table 1) [6]. In non-anophelines, the sylvatic African subpopulation of *Aedes aegypti formosus* differs genetically, morphologically and ecologically from the sympatric domestic form, *Aedes aegypti aegypti*, to an extent greater than the differences among the worldwide populations of the domestic form [21]. In *C. pipiens*, a comparable pattern was demonstrated for sympatric commensal-urban and wildtype-rural ecotypes, between which the genetic differences are low ($F_{ST}$=0.06), but behavioural, ecological and physiological differences are substantial [22]. These examples of strong heterogeneity between populations, mostly revealed by non-molecular markers under selection, expose the limitation of geographical sampling as the ultimate basis for understanding species population structure. Studies of ecology, gene-expression and mating patterns in addition to the standard molecular genetic approach are required. The studies will help define the extent of reproductive isolation resulting from these discontinuities, their role in speciation and whether these dimensions truly represent domesticity (as proposed for *C. pipiens* [22] and *An. funestus* [17]), or aridity in the case of *An. gambiae* chromosomal forms. Furthermore, the low levels of population differentiation revealed by molecular genetic techniques suggest that this approach is an inefficient way for detecting inter-population differences in malaria transmission rates. Correlations have been observed between chromosomal inversions and heterogeneities in malaria transmission [23]. Such differences in transmission, if substantiated by further study, could be determined by a small number of genes within the inversions, where suppression of recombination would preserve particular allele complements.

## Relating population structure to contemporary gene flow

The motivation for studies of the population structure of malaria vectors was to estimate gene flow in the expectation that this would allow a prediction of the rate of spread of genes of interest (e.g. genes conferring insecticide resistance) between populations. Whereas a robust description of population structure (in terms of genetic similarity among populations) is reasonably achieved by differentiation indices such as $F_{ST}$ [24], the interpretation of the structure in terms of gene flow or phylogenetic relationships between populations is currently obscured.

### *There are few barriers to gene flow*

The low level of population differentiation observed in primary vectors is a reflection, at least in part, of the relatively few physical barriers to gene flow. The large differentiation among island populations of *An. arabiensis* [25] compared with continental populations [2,26,27] suggest that wide stretches of ocean restrict genetic exchange, but founder effects on the islands will also decrease the estimates of gene flow. Non-marine barriers were rarely implicated, although mountainous regions probably restricted gene flow in *An. albimanus* in Central America [28] and in *Anopheles punctipennis* in North America [29]. The arid valleys and cool highlands of the Rift Valley Complex restricted gene flow between western and eastern Kenyan populations of *An. gambiae* but not of *An. arabiensis*, although a bottleneck in the eastern populations also increased differentiation and inflated the effect of the Rift Valley Complex [10,27,30,31].

### *Differences in effective population size and recent demographic expansion bias estimates of gene flow*

Differences in the effective population size ($N_e$) among populations will account for a non-negligible part of genetic differentiation [24]. Significant heterogeneity in $N_e$ exists between populations within the *An. gambiae* complex [10,25,26,30,32,33]. Lehmann *et al*. [10] demonstrated that this heterogeneity within the *An. gambiae* of the savanna cytotype was region specific and did not result from population-specific characteristics or sampling error. Within *An. arabiensis* marginal populations (from the edges of the species range on the African continent [26] or from the Mascarene islands in the Indian Ocean [25]) presented significantly lower genetic diversity than samples from the core of the species range. Similar patterns were observed in *An. darlingi* where diversity indices were low in a population from Belize that was separated from the main species range in South and south-Central America [6], and in *An. pseudopunctipennis* where mean heterozygosity was lower in a population from Grenada [7]. Interpopulation differences in $N_e$ were also suggested for *An. albimanus* in Panama based on the mtDNA haplotype diversity – a variation that could be related to the extensive mosquito control programmes in the region [8]. The heterogeneity in $N_e$ seems to involve mainly populations in marginal habitats, isolated from the core species range or those subject to intensive control.

Unstable historical demography, in particular expansions, has been repeatedly discussed as a major factor shaping population structure [1,2,11,34]. In practical terms, this has been ignored as evidenced by the focus on estimates of gene flow derived from

differentiation. mtDNA studies showed evidence for population instability in the *An. gambiae* and *An. dirus* complexes [11,34], although both studies noted that this might reflect the locus-specific effect of a haplotype sweeping to fixation rather than a demographic expansion. These initial findings were confirmed by multi-locus tests that detect departures from mutation-drift-migration equilibrium. Unstable demography has been demonstrated in *An. gambiae, An. arabiensis* and *An. dirus* (A and D), resulting in disequilibrium between migration and drift, thereby preventing accurate estimation of levels of gene flow ($N_m$) from $F_{ST}$ estimates [4,31].

*Recent speciation or introgression*

The shallowness of the structure in anophelines is illustrated by the failure of molecular approaches to produce robust phylogenies within species complexes. In general, the markers used in population studies have either proved insufficiently variable (e.g. the internal transcribed spacer region of the rDNA [16]) or when variable methodological issues complicate their use (e.g. microsatellites [4]). mtDNA has been the marker of choice for resolving phylogenetic relationships between species and within species complexes [34–36]. However, the number of mutational steps that separate haplotypes is few, and fixed differences are rarely found even between species. Therefore, the major problem confronting phylogeny-based approaches in *An. gambiae*, *An. funestus* and *An. dirus* complexes, and probably by extension in *Anopheles culicifacies* and *Anopheles stephensi* is that the species have probably only separated recently and have not experienced the $4N_e$ generations of isolation necessary for evolving separate, mutually-exclusive mitochondrial lineages [35,37]. Presently, it has not been possible to distinguish conclusively between the conflicting hypotheses of retention of ancestral haplotypes (i.e. haplotypes that were found in the ancestor of both species) and ongoing introgression. Introgression, the favored hypothesis, is supported by laboratory experiments, which demonstrated that sections of nuclear DNA can introgress between species [38].

*Resolving the status of taxa within* An. gambiae s.s.

Studying chromosomal inversions in *An. gambiae* (Mopti) populations from Mali, Touré *et al*. [39] described a dynamic system in which different inversion arrangements change in frequency, apparently in response to climatic fluctuations. Highly significant differences were even documented between collections from the same sampling site, separated by less than one generation (i.e. 10–14 days). As a result of the correlation between inversions and, seasonal and spatial ecoclimatic variations [15], debate has centred on the relative influences of reproductive isolation and natural selection in maintaining the heterokaryotype deficits observed between the five chromosomally recognized units. Recently, a series of studies found that X-linked rDNA could be used to distinguish two widespread molecular types termed M and S [40,41], although most autosomal markers did not support these data (see Ref. [42] and references therein). In Mali and Burkina Faso, the M form corresponds to the Mopti chromosomal form, whereas sympatric populations of Savanna and Bamako belong to the S molecular form. However, the correspondence between chromosomal and molecular forms does not hold true elsewhere in West Africa [41,43].

Tripet *et al.* [44] characterized the molecular type of individual females and the sperm with which they had been inseminated, and thereby demonstrated strong ASSORTATIVE MATING within the S and M forms. It is thought that the M and S molecular forms correspond to the Mopti and Bamako chromosomal forms, respectively, because these are the only two forms that have been observed in the area [44]. Mating between molecular forms occurred at a frequency of 1% and there is likely to be considerable gene flow between the M and S forms because there is no evidence for post-zygotic selection, although this still requires further investigation. What remains to be resolved is whether this is a stable situation, or if the M and S forms are in a process of incipient speciation [45]. What maintains associative mating within these forms is unknown. The differences in the cuticular hydrocarbons observed between chromosomal forms, which are strongest in sympatric populations and fail to distinguish ALLOPATRIC forms, suggest that these hydrocarbons play a role in a mate-recognition systems as observed in some *Drosophila* species [45–47]. Selection maintains pre-zygotic mating barriers more strongly in sympatric populations than between contiguous allopatric populations [47]. This would explain why the linkage between rDNA molecular type and chromosomal form breaks down in areas outside Mali and Burkina Faso, and indicates that the mate-recognition system is not associated with the 2R inversions [15]. Furthermore this suggests that inversion arrangements are sometimes coincidental but do not define certain biological entities within *An. gambiae*. Currently, the molecular M and S designation might more accurately characterize the taxa within *An. gambiae*. Therefore, whereas gene flow might be restricted by mate recognition systems which the rDNA data suggests are associated with the X chromosome genetic exchange is sufficient to maintain homogeneity on loci outside regions that are differentially selected.

*What happens to mosquitoes during the dry season?*

Some of the studies that revealed differences in effective size between populations also permitted investigation of one of the great unknowns of malaria entomology, namely the dry-season survival of vector populations in the dry savannas of Africa. The direct (ecological) approach to study the dry-season

dynamics using mark–release–recapture experiments [48] is a challenging one particularly when mosquitoes cannot be found or are very scarce [49,50]. By sampling a population repeatedly, with samples separated by a number of generations, it is possible to quantify the temporal changes in allele frequencies at neutral loci and thereby estimate $N_e$ [51]. This technique has been used to estimate the effective population size of *An. gambiae* in Kenya [52] and *An. arabiensis* in West Africa [32,33]. Reported estimates of $N_e$ were in the thousands for both *An. gambiae* and *An. arabiensis*. Such estimates provide convincing evidence against recurrent extinction and recolonization events or dry-season bottlenecks, suggesting that vector populations are maintained locally throughout the year despite the extreme variations in abundance, that are observed in the field. The hypothesis of a diffused DEME [33,52] was proposed to reconcile both ecological and genetic data [50]. According to this hypothesis, a small number of mosquitoes are maintained locally, perhaps through aestivation in the absence of suitable breeding sites, and the deme is spread across a large geographic area (between 20 and 50 km in diameter). This allows a relatively high, effective population size to be maintained despite low densities of vectors, which are often below sampling sensitivity. These genetic studies did not resolve the question of how do mosquitoes survive the dry season in Africa, but they did provide clear evidence against recurrent founder effects by a limited number of individuals. These studies also highlighted that mosquitoes from the *An. gambiae* complex are capable of extensive mobility at a local geographical scale (≤50 km). Hence, in order to reduce vector numbers, control activities need to be spread over a wide geographical area. Isolated interventions at the village level will fail because of an influx of mosquitoes from neighbouring untreated areas. Furthermore, this suggests that hypothetical source populations, acting as reservoirs of mosquitoes during the dry season, do not exist in these regions, and targeted control of geographically confined sites is unlikely to be successful. This can be illustrated by the failure of insecticide and sterile-insect release based control campaigns, which resulted from a combination of operational failures and large vector population sizes. However, the results described are from rural environments, and there is evidence that the situation could be different in urban settings, where malaria vectors might be more spatially structured towards the periphery of densely populated areas [53].

Reconciling these indirect estimates of effective population size and gene flow with ecological estimates of total population size and rates of migration is problematic because each method relies on a number of unknown parameters. Therefore, even if both the genetic and ecological estimates are correct and exact [49], they are likely to disagree without either being wrong. Ecological estimates of dispersal and population size based on mark–release–recapture aim to measure the average distance travelled by individuals and the average number of adults present during a defined period in a geographical area related to the distance travelled during this time. By contrast, genetic estimates aim to measure extreme values of distance travelled during the life span of the organism (coupled with successful breeding), and the extreme low values in the number of breeding adults in a series of generations in a deme (with adjustments for variance in progeny numbers). The distributions of distance travelled and absolute population size per generation will have tails long away enough from the mean to generate substantial discrepancy. The discrepancy will only increase with heterogeneity in time and space (e.g. long distance migration occurring only with special weather conditions). Thus, whereas the ecological and genetic approaches are complementary, a prerequisite for relating them is knowledge of the underlying distributions. Because these distributions are not known for any anopheline, the efforts to reconcile genetical and ecological estimates require subjective adjustments, which are difficult to justify.

## Conclusions

Contemporary gene flow and drift are not the primary forces shaping some of the more important population divisions of anopheline vectors. In addition, the methods to estimate gene flow are sensitive to departures from mutation-drift equilibrium, variation in $N_e$ and the differing histories of populations, all of which were documented in anophelines. Therefore, the presently available estimates of gene flow could be misleading (see [54] for discussion). Nevertheless, knowledge of the population structure of anopheline vectors has been gained and, to an extent, its underlying causes, which advances our understanding of the past, present and future of these species. The rapid development of analytic approaches will be the key to the separation of contemporary gene flow from other causes.

The limitations of large-scale geographical studies for detecting and understanding heterogeneities in the population structure of these insects need to be recognized. The forces that cause homogeneity across large distances are overcome by selection pressure operating at often-overlooked scales. Ecological studies on dispersal, population dynamics of adults and larvae, mating behaviour, and phenotypic adaptations to various conditions are the key for detecting and understanding the functional processes shaping population structure. Therefore, it is anticipated that studies in areas where distinct populations exist in sympatry or where selection pressures, such as aridity and insecticide use, are defined will be the most valuable in resolving these problems.

**Acknowledgements**
We would like to thank Philip McCall, Joao Pinto, Francois Rousset, Harold Townson and Steven Sinkins and the two anonymous referees for valuable comments on the manuscript, and Greg Lanzaro and Cathy Walton for providing prepublication copies of papers.

**References**

1 Lehmann, T. *et al*. (1996) Genetic differentiation of *Anopheles gambiae* populations from East and West Africa: comparison of microsatellite and allozyme loci. *Heredity* 77, 192–200
2 Donnelly, M.J. and Townson, H. (2000) Evidence for extensive genetic differentiation among populations of the malaria vector *Anopheles arabiensis* in eastern Africa. *Insect Mol. Biol.* 9, 357–367
3 Besansky, N.J. (1999) Complexities in the analysis of cryptic taxa within the genus *Anopheles*. *Parassitologia* 41, 97–100
4 Walton, C. *et al*. (2001) Genetic population structure in *Anopheles dirus* mosquitoes in southeast Asia. *Mol. Ecol.* 10, 569–580
5 Rongnoparut, P. *et al*. (1999) Estimates of gene flow among *Anopheles maculatus* populations in Thailand using microsatellite analysis. *Am. J. Trop. Med. Hyg.* 60, 508–515
6 Manguin, S. *et al*. (1999) Population structure of the primary malaria vector in South America *Anopheles darlingi*, using isozyme, random amplified polymorphic DNA, internal transcribed spacer 2, and morphologic markers. *Am. J. Trop. Med. Hyg.* 60, 364–376
7 Manguin, S. *et al*. (1995) Biochemical systematics and population genetic structure of *Anopheles pseudopunctipennis*, vector of malaria in Central and South America. *Am. J. Trop. Med. Hyg.* 53, 362–377
8 De Merida, A.M. *et al*. (1999) Mitochondrial DNA variation among *Anopheles albimanus* populations. *Am. J. Trop. Med. Hyg.* 61, 230–239
9 Lounibos, L.P. and Conn, J.E. (2000) Malaria vector heterogeneity in South America. *Am. Entomol.* 46, 238–249
10 Lehmann, T. *et al*. (1999) The Rift Valley complex as a barrier to gene flow for *Anopheles gambiae* in Kenya. *J. Hered.* 90, 613–621
11 Besansky, N.J. *et al*. (1997) Patterns of mitochondrial variation within and between African malaria vectors, *Anopheles gambiae* and *An. arabiensis*, suggest extensive gene flow. *Genetics* 147, 1817–1828
12 Tabachnick, W.J. (1991) The yellow fever mosquito. *Am. Entomol.* Spring, 16–24
13 Chevillon, C. *et al*. (1995) Population-structure and dynamics of selected genes in the mosquito *Culex pipiens*. *Evolution* 49, 997–1007
14 Coluzzi, M. *et al*. (1979) Chromosomal differentiation and adaptation to human environments in the *An. gambiae* complex. *Trans. R. Soc. Trop. Med. Hyg.* 73, 483–497
15 Touré, Y.T. *et al*. (1998) The distribution and inversion polymorphism of chromosomally recognised taxa of the *Anopheles gambiae* complex in Mali, West Africa. *Parassitologia* 40, 477–511
16 Mukabayire, O. *et al*. (1999) Mitochondrial and ribosomal internal transcribed spacer (ITS2) diversity of the African malaria vector *Anopheles funestus*. *Mol. Ecol.* 8, 289–297
17 Dia, I. *et al*. (2000) Spatial and temporal variations of the chromosomal inversion polymorphism of *Anopheles funestus* in Senegal. *Parasite* 7, 179–184
18 Costantini, C. *et al*. (1999) Chromosomal and bionomic heterogeneities suggest incipient speciation in *Anopheles funestus* from Burkina Faso. *Parassitologia* 41, 595–611
19 Kreutzer, R.D. *et al*. (1972) Inversion polymorphism in the salivary gland chromosomes of *Anopheles darlingi*. *Mosq. News* 32, 555–565
20 Charlwood, J.D. (1996) Biological variations in *Anopheles darlingi* Root. *Mem. Inst. Oswaldo Cruz* 91, 391–398
21 Powell, J.R. *et al*. (1980) Genetics and the origin of a vector population: *Aedes aegypti*, a case study. *Science* 208, 1385–1387
22 Chevillon, C. *et al*. (1995) Commensalism, adaptation and gene flow: mosquitoes of the *Culex pipiens* complex in different habitats. *Genet. Res.* 66, 147–157
23 Petrarca, V. and Beier, J.C. (1992) Intraspecific chromosomal polymorphism in the *Anopheles gambiae* complex as a factor affecting malaria transmission in the Kisumu area of Kenya. *Am. J. Trop. Med. Hyg.* 46, 229–237
24 Wright, S. (1978). *Evolution and the genetics of populations. Volume IV. Variability among and within populations* (2nd edn), University of Chicago Press
25 Simard, F. *et al*. (1999) High amounts of genetic differentiation between populations of the malaria vector *Anopheles arabiensis* from West Africa and eastern outer islands. *Am. J. Trop. Med. Hyg.* 60, 1000–1009
26 Donnelly, M.J. *et al*. (1999) Population structure in the malaria vector, *Anopheles arabiensis* Patton, in East Africa. *Heredity* 83, 408–417
27 Kamau, L. *et al*. (1999) Analysis of genetic variability in *Anopheles arabiensis* and *Anopheles gambiae* using microsatellite loci. *Insect Mol. Biol.* 8, 287–297
28 De Merida, A.M. *et al*. (1995) Variation in ribosomal DNA intergenic spacers among populations of *Anopheles albimanus* in South and Central America. *Am. J. Trop. Med. Hyg.* 53, 469–477
29 Fairley, T.L. *et al*. (2000) Effects of local geographic barriers and latitude on population structure in *Anopheles punctipennis* (Diptera: Culicidae). *J. Med. Entomol.* 37, 754–760
30 Lehmann, T. *et al*. (2000) The Rift Valley complex as a barrier to gene flow for *Anopheles gambiae* in Kenya. The mtDNA perspective. *J. Hered.* 91, 165–168
31 Donnelly, M.J. *et al*. (2001) Evidence for recent population expansion in the evolutionary history of the malaria vectors *Anopheles arabiensis* and *Anopheles gambiae*. *Mol. Biol. Evol.* 18, 1353–1364
32 Taylor, C.E. *et al*. (1993) Effective population size and persistence of *Anopheles arabiensis* during the dry season in West Africa. *Med. Vet. Entomol.* 7, 351–357
33 Simard, F. *et al*. (2000) Persistence of *Anopheles arabiensis* during the severe dry season conditions in Senegal: an indirect approach using microsatellite loci. *Insect Mol. Biol.* 9, 467–479
34 Walton, C. *et al*. (2000) Population structure and population history of *Anopheles dirus* mosquitoes in southeast Asia. *Mol. Biol. Evol.* 17, 962–974
35 Besansky, N.J. *et al*. (1994) Molecular phylogeny of the *Anopheles gambiae* complex suggests genetic introgression between principal malaria vectors. *Proc. Natl. Acad. Sci. U. S. A.* 91, 6885–6888
36 Conn, J.E. *et al*. (1997) Mitochondrial DNA variation within and between two species of neotropical anopheline mosquitoes (Diptera:Culicidae). *J. Hered.* 88, 98–107
37 Avise, J.C. *et al*. (1984) Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.* 20, 99–105
38 della Torre, A. *et al*. (1997) Selective introgression of paracentric inversions between two sibling species of the *Anopheles gambiae* complex. *Genetics* 146, 239–244
39 Touré, Y.T. *et al*. (1994) Ecological genetic studies in the chromosomal form Mopti of *Anopheles gambiae s.str*. in Mali, west Africa. *Genetica* 94, 213–223
40 Favia, G. *et al*. (1997) Molecular identification of sympatric chromosomal forms of *Anopheles gambiae* and further evidence of their reproductive isolation. *Insect Mol. Biol.* 6, 377–383
41 della Torre, A. *et al*. (2001) Molecular evidence for incipient speciation within *Anopheles gambiae s.s.* in West Africa. *Insect Mol. Biol.* 10, 9–18
42 Black, W.C., IV and 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
43 Gentile, G. *et al*. (2001) Attempts to molecularly distinguish cryptic taxa in *Anopheles gambiae s.s. Insect Mol. Biol.* 10, 25–32
44 Tripet, F. *et al*. (2001) DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. *Mol. Ecol.* 10, 1725–1732
45 Powell, J.R. *et al*. (1999) Population structure, speciation and introgression in the *Anopheles gambiae* complex. *Parassitologia* 41, 101–114
46 Milligan, P.J. *et al*. (1993) A study of the use of gas chromatography of cuticular hydrocarbons for identifying members of the *Anopheles gambiae* (Diptera:Culicidae) complex. *Bull. Entomol. Res.* 83, 613–624
47 Higgie, M. *et al*. (2000) Natural selection and the reinforcement of mate recognition. *Science* 290, 519–521
48 Service, M.W. (1993) *Mosquito Ecology: Field Sampling Methods* Elsevier Science
49 Touré, Y.T. *et al*. (1998) Mark-release-recapture experiments with *Anopheles gambiae s.l.* in Banambani Village, Mali, to determine population size and structure. *Med. Vet. Entomol.* 12, 74–83
50 Taylor, C. *et al*. (2001) Gene flow among populations of the malaria vector, *Anopheles gambiae*, in Mali, West Africa. *Genetics* 157, 743–750
51 Waples, R.S. (1991) Genetic methods for estimating the effective population size of cetacean populations. In *Genetic Ecology of Whales and Dolphins* (Hoelzel, A.R., ed.) pp. 279–300, International Whaling Commission (Special Issue No. 13). Cambridge University Press
52 Lehmann, T. *et al*. (1998) The effective population size of *Anopheles gambiae* in Kenya: implications for population structure. *Mol. Biol. Evol.* 15, 264–276
53 Trape, J-F. *et al*. (1992) Vector density gradients and the epidemiology of urban malaria in Dakar, Senegal. *Am. J. Trop. Med. Hyg.* 47, 181–189
54 Whitlock, M.C. and McCauley, D.E. (1999) Indirect measures of gene flow and migration: $F_{ST}$ not equal to $1/(4N_m+1)$. *Heredity* 82, 117–125
55 Miles, S.J. (1978) Enzyme variation in the *Anopheles gambiae* Giles group of species (Diptera:Culicidae). *Bull. Ent. Res.* 68, 85–96
56 Schaffner, F. *et al*. (2000) Genetic differentiation of *Anopheles claviger s.s.* in France and neighbouring countries. *Med. Vet. Entomol.* 14, 264–271
57 Conn, J.E. *et al*. (1999) Molecular population genetics of the primary neotropical malaria vector *Anopheles darlingi* using mtDNA. *J. Am. Mosq. Control. Assoc.* 15, 468–474